

## DETECTION OF MUTATIONS IN GENES BY SPECIFIC LNA PRIMERS.

### FIELD OF THE INVENTION

The present invention comprises a method for detecting variant nucleic acids, a set of oligonucleotides suitable for this purpose, a reagent kit for performing the method of the invention and various uses and applications of the method in identifying specific gene sequences and diagnosing diseases and infections.

### BACKGROUND OF THE INVENTION

Today, the testing of samples for the presence of certain nucleic acids in nucleic acid sequences gains increasingly more importance. This is partly due to the fact that the nucleotide sequence of a nucleic acid is a unique feature of each organism. First, a number of diseases are genetic in the sense that the nucleotide sequence for a "normal" gene is in some manner changed. Such a change could arise by the substitution of one base for another. Changes comprising more than one base can also be perceived. Given that three bases code for a single amino acid, a change in one base (a point mutation) could result in a change in the amino acid which, in turn, could result in a defective protein being made in a cell. Sickle cell anemia is a classic example of such a genetic defect caused by a change of a single base in a single gene; the beta-globin gene (A→T transversion at codon 6). Important point mutations can also be found for example in LDL receptors (Atherosclerosis (1992), 96: 91-107) or in the apolipoprotein-B-gene (CGG→CAG mutation of codon 3500 (Proc. Natl. Acad. Sci. U.S.A. (1989) 86:587-591). Other examples of diseases caused by single gene defects include Factor IX and Factor VIII deficiency, breast cancer, cystic fibrosis, Factor V Leiden, Fragile X, Huntington disease, myotonic dystrophy, Haemophilia A, Haemophilia B, Neurofibromatosis type I, adenosine deaminase deficiency, purine nucleotide phosphorylase deficiency, ornithine transcarbamylase deficiency, argininosuccinate synthetase deficiency, beta-thalassemia, alpha-1 antitrypsin deficiency, glucocerebrosidase deficiency, phenylalanine hydroxylase deficiency and hypoxanthine-guanine phosphoribosyltransferase deficiency.

Second, the main cause of cancer is considered to be alterations in the cellular genes which directly or indirectly control cell growth and differentiation. There are at least thirty families of genes, called oncogenes, which are implicated in human tumor formation. Members of one such family, the *ras* gene family, are frequently found to be mutated in

human tumors. In their normal state, proteins produced by the *ras* genes are thought to be involved in normal cell growth and maturation. Mutation of the *ras* gene, causing an amino acid alteration at one of three critical positions in the protein product, results in conversion to a form which is implicated in tumor formation. A gene having such a mutation is said to be "mutant" or "activated." Unmutated *ras* is called "wild-type" or "normal" *ras*. It is thought that such a point mutation leading to *ras* activation can be induced by carcinogens or other environmental factors. Over 90% of pancreatic adenocarcinomas, about 50% of adenomas and adenocarcinomas of the colon, about 50% of adenocarcinomas of the lung and carcinomas of the thyroid, and a large fraction of haematological malignancies such as acute myeloid leukemia, lymphomas and myelodysplastic syndrome have been found to contain activated *ras* oncogenes. Overall, some 10 to 20% of human tumors have a mutation in one of the three *ras* genes (*H-ras*, *Ki-ras*, or *N-ras*).

Another example of a gene which is highly involved in the development of cancer is the TP53 gene. It is altered by mutations and/or deletions in more than half of all human cancers. The point mutations are scattered over more than 250 codons and mostly occur as missense mutations. In this respect, the TP53 gene differs from other tumor suppressor genes such as the retinoblastoma tumor suppressor gene (*Rb1*) and the p16 gene which are most frequently inactivated by deletions or nonsense mutations.

Most malignant tumors show alterations in both alleles of the TP53 gene. This usually involves the complete deletion of one allele and inactivation of the other allele by missense mutations. The result is either a complete lack of TP53 protein or expression of an altered protein. The missense mutations in the highly conserved regions of TP53 have also been associated with increased level of TP53 protein. This seems to result from mutation induced conformational changes, which stabilize the protein and extends its half-life from 4 to 8 hours. The majority of missense mutations cluster in the four highly conserved domains in the central core of the protein. This region is responsible for the sequence specific DNA binding and is therefore of critical importance for the functional integrity of TP53. Seven mutational hot spots have been identified within these domains. These are located at amino acid residues 175, 213, 245, 248, 249, 273, and 282.

Third, infectious diseases are caused by parasites, micro-organisms and viruses all of which have their own nucleic acids. The presence of these organisms in a sample of biological material is often determined by a number of traditional methods (e.g., culture).

Each organism has its own unique genome and if there are genes or sequences of nucleic acids that are specific to a single species (to several related species, to a genus or to a higher level of relationship), this sequence will provide a "fingerprint" for that organism (or species, etc.), e.g. in the gene of the reverse transcriptase of the HIV virus (A→T mutation in codon 215: Science (1989) 246:1155-1158). Examples of other viruses include HPV, EBV, HSV, Hepatitis B and C and CMV. Examples of micro-organisms include bacteria and more particularly include various strains of mycoplasma, legionella, mycobacteria, chlamydia, candida, gonocci, shigella and salmonella. As information on the genomes from more organisms are obtained by the scientific community the repertoire of micro-organisms that can be identified by the present invention will increase. In the nucleic acids of some bacterial strains, a particularly great similarity is found in the sequence of their ribosomal genes and their rRNA.

Current attempts in the field of examining samples for different nucleotide sequences focus on the use of only one single difference in the sequence of nucleotides in order to be able to discriminate between nucleic acids. Such differences may be a consequence of e.g. nucleotide exchanges caused by point mutation or in the case of micro-organisms a consequence of inter-species differences. Natural examples of such closely related nucleic acids are alleles, i.e. alternative variants of sequences of a given gene on a defined site on a chromosome.

In each example set forth above one can isolate nucleic acids from a sample and determine if the sample contains any of the above mentioned sequences, i.e. sequences specific for "genetic disease", cancer, infectious diseases or infectious organisms, by identifying one or more sequences that are specific for a diseases or organism. A difficulty when identifying these differences or changes in the nucleotide sequence is however that the detection is not readily applicable in those instances where the number of copies of the target sequence present in a sample is low. In such instances it is difficult to distinguish signal from noise. One way around this problem is to increase the signal. Accordingly, a number of methods have been described to amplify the target sequences present in a sample. One of the best known and widely used amplification methods is the polymerase chain reaction (referred to as PCR) which is described in detail in US 4,683,195, US 4,683,202 and US 4,800,159.

Based upon the PCR technique a number of methods for detection of sequence variations have been described. From Oncogene Research 1 (1989), 235-241 and Nucl. Acids Res. 17 (1989), 8093-8099 a method is known where the area which presumably contains the allelic variant is first amplified in a PCR using specially designed primers and is then

5 treated with a restriction enzyme. The alleles can then be diagnosed once they have been analyzed with restriction fragment length polymorphisms (RFLP). Electrophoretic separation of the cleavage products according to size then reveals whether the corresponding allele was or was not contained in the probe. The disadvantage of this procedure is that it requires specific restrictive digestion. Apart from the fact that this is a cumbersome procedure for each mutation that does not already produce an RFLP it is necessary that a  
10 primer can be designed to be adjacent to the point mutation which should allow digestion with a restriction enzyme that cleaves exactly at the given site. This may be difficult due to the reasons listed in said publications.

15 EP-A-0 332 435 and US 5,605,794 describe a method for selectively detecting a nucleic acid which differs from an adjacent nucleic acid by only one nucleotide. The effect employed here is the following one: from the oligonucleotides which are hybridized to the nucleic acid to be detected only those can be theoretically extended by means of enzymes where the one nucleotide, which is terminal in direction of extension, is complementary to  
20 the corresponding nucleotide of the nucleic acid (of the one allele) to be detected. The oligonucleotide is hence selected such that it is only complementary to the nucleic acid to be tested. Thus, the oligonucleotide hybridized to the other allele is theoretically not extended. It turned out, however, that in practice the oligonucleotide hybridized to the other allele is, though only to a minor extent, also extended. This reduces the sensitivity and,  
25 particularly, the specificity of the method. Non-specific extension may easily occur especially when T is part of the 3'-terminal mismatch or when the mismatch is a C:A mismatch (Kwok et al. (1990). Nucleic Acids Research, 18:999-1005). In order to increase specificity, EP-A-0 332 435 proposes to select the nucleotide sequence of an oligonucleotide such that the terminal area contains another nucleotide that is not complementary to the  
30 corresponding nucleotide of the two nucleic acids. For the detection of both alleles two reactions must be carried out with only one of two alleles being detected per reaction. This procedure requires the synthesis of two allele-specific primers and one complementary strand primer. The sample is amplified in two reactions: once in a PCR with the primer of the complementary strand and one of the allele-specific primers, and in the sec-  
35 ond parallel reaction, a PCR, it is amplified with the complementary strand primer and the

second allele-specific primer. If the suspected allele-specific PCR product is not detected in one of the reactions, it is assumed that the respective allele is not present in the sample. Since homozygous DNA-samples contain only one of the two alleles which can be detected in only one of the two reactions, it is necessary to use two additional primers  
 5 which produce the same control product in all reactions. This control product is different from the specific product in order to control the efficiency of the respective PCR of the other allele and to establish the absence of the respective allele. If a control product is present in the PCR product but no specific product of the allele is found, the sample is not likely to contain the allele tested for in the reaction. In this method, the presence or ab-  
 10 sence of two alleles must be established in two separate reactions and each individual PCR must comprise a control PCR.

Biochem. Biophys. Res. Commun. (1989), 160:441-447 proposes to increase the selectivity by decreasing the dNTP concentration. Even if this additional measure is taken, the  
 15 detection of alleles in separate batches can yield non-specific products.

In a ligase chain reaction (WO 89/09835), thermostable ligase is used to specifically link two adjacent oligonucleotides. This occurs only if they are hybridized to a complementary target at a stringent hybridization temperature and if base-pairing at the site of linkage is  
 20 complete. If two alleles differ from each other as a consequence of a mutation at the linkage site, the above condition of complete base-pairing is fulfilled for only one of the alleles. Two additional oligonucleotides, which are complementary to the first two, are then necessary to amplify the ligation product in the ligase chain reaction. To date, the detection of two alleles requires two reactions with at least six oligonucleotides, and the amplifi-  
 25 cation product is detected with a radioactive label (Proc. Natl. Acad. Sci. U.S.A. (1991), 88:189-193).

From Proc. Natl. Acad. Sci. U.S.A. (1985), 82:1585-1588 and from New England Journal of Medicine (1987), 317:985 a method of detecting alleles is known which is based upon  
 30 differential hybridization of "allele-specific" oligonucleotides (ASO) with the alleles under examination. Two oligonucleotides, each 20 bp in length, for example, are synthesized. Each matches one of the two different alleles but has a mismatch to the other allele located in the middle of the oligonucleotide sequence. Discrimination between alleles is then possible by differential hybridization with labelled oligonucleotides. This applies to  
 35 the analysis of both human genomic DNA and PCR products. Direct and discrete analysis

of genomic DNA is also possible with this method but requires additional digestion and electrophoresis.

Nucleic Acids Research (1989), 17:2437-2448 and EP-A-0 333 465 describe a method of testing pre-amplified human genomic DNA for the presence of various alleles in a few additional PCR cycles by competition of allele-specific primers (competitive oligonucleotide priming = COP). The above described ASO-technique is then converted into a PCR technique. In the original ASO-technique, an error rate of 5% caused by cross hybridization is acceptable since a comparison of the signal intensities during corresponding controls allows an unequivocal interpretation of the results. In a PCR reaction, however, where the primers are allele-specific oligonucleotides, the error rate for a sample that contains only one of the alleles would after ten cycles amount to 12% if this error occurred in a reagent mixture where both alleles are amplified. It could indeed be demonstrated that primer competition increased selectivity, however, the area of interest of the genomic DNA was first amplified in a PCR and the analysis for the different alleles was then carried out in ten subsequent cycles. Two allele-specific primers and a complementary strand primer were used in these cycles and in two reactions one of the allele-specific primers was radioactively labelled. A selective detection of different alleles has been demonstrated for oligonucleotides of 12 to 16 bases in length whereas longer oligonucleotides under the given conditions also yielded non-specific products.

There is a need for a simple and still very specific method for directly detecting at least one single base difference in a sample containing nucleic acids such as genomic DNA in which detection steps are minimized resulting in a method which may be performed quickly, accurately and easily with minimal operator skills.

Methods that are based on differential hybridization can only be applied in certain situations and are moreover very complex and susceptible to interference. Also, pre-amplification is a procedural step, e.g. during COP, which preferably is eliminated.

30

All of the above mentioned methods for detecting variant nucleic acids have the same feature of relying on unmodified nucleotides as the discriminating factor in the detection of the variant nucleic acids.

## SUMMARY OF THE INVENTION

The present invention provides a simple method for the specific detection of variant nucleic acids which does not require digestion by restriction enzymes, allows amplifying the nucleic acids in one reagent mixture and does not depend on subsequent enzymatic reactions. A crucial component in the invention is LNA which is a novel class of DNA analogues that possess some extraordinary features that make it a prime candidate for improving *in vitro* DNA diagnostics. The LNA monomers are bi-cyclic compounds structurally very similar to RNA-monomers, see formula I. LNA share most of the chemical properties of DNA and RNA, it is water-soluble, can be separated by gel electrophoreses, ethanol precipitated etc (Tetrahedron, 54, 3607-3630 (1998)). However, introduction of LNA monomers into either DNA or RNA oligos results in unprecedented high thermal stability of duplexes with complementary DNA or RNA, while, at the same time obeying the Watson-Crick base-pairing rules. This high thermal stability of the duplexes formed with LNA oligomers together with the finding that primers containing 3' located LNA(s) are substrates for enzymatic extensions, e.g. the PCR reaction, is used in the present invention to significantly increase the specificity of detection of variant nucleic acids in the *in vitro* assays described in the application. The amplification processes of individual alleles occur highly discriminative (cross reactions are not visible) and several reactions may take place in the same vessel.

## DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a method of detecting variant nucleic acids having nucleotide sequences which differ from one another in at least one position A. The method comprises at least two steps:

25

a) Performing an extension reaction with at least one diagnostic oligonucleotide as a primer for the extension reaction, the at least one diagnostic oligonucleotide being a LNA-containing oligonucleotide wherein the LNA(s) is/are preferably found at position A or at neighbouring positions to position A, most preferably at the 3'-end of the oligonucleotide,

30

b) detecting the extension product, e.g. by standard gel electrophoresis, capillary electrophoresis or various hybridizations assays.

Other embodiments of the invention are further extension and amplification using the at least one oligonucleotide, a set of oligonucleotides, different sets of oligonucleotides, a reagent kit and various applications and uses of said method.

- 5 The "target nucleic acids" in accordance with the invention are nucleic acids that may be contained in a sample and the presence of which is of interest, the nucleotide sequences of the target nucleic acids (Q' and Q) are substantially identical, but the nucleic acid Q' differs in at least one position of its nucleotide sequence from the nucleotide sequence of nucleic acid Q. The position of the difference in the nucleotide sequence is referred to as
- 10 "position A". Position A is also referred to as a polymorphous site and the presence of a position A in a nucleic acid is referred to as a gene polymorphism. If the variant nucleotide sequences comprise several different nucleic acids (Q'<sub>1</sub>, Q'<sub>2</sub>, . . . Q'<sub>e</sub>) the positions of differences will be referred to as A<sub>1</sub>, A<sub>2</sub>, . . . A<sub>e</sub>. Such differences may occur, for example, as a consequence of point mutations or may be caused by deletions or insertions of one or
- 15 several bases. The nucleic acids to be detected can be RNA or DNA and especially in the diagnosis of bacteria rRNA has proven to be particularly suitable.

- The term "diagnostic oligonucleotide" refers to an oligonucleotide, which at position A has a nucleotide that is complementary to the nucleotide found at position A of the target nucleic acid to be detected and which is capable of discriminating between different target
- 20 nucleic acid sequences (Q, Q'<sub>1</sub>, Q'<sub>2</sub>, . . . Q'<sub>e</sub>) under the hybridization conditions given and only initiate extension of the target nucleic acid sequences which contain at a position A a nucleotide which is complementary to the nucleotide at position A of the diagnostic oligonucleotide.

25

The term "product target" refers to those nucleotides that are extension products of the original target nucleic acids. Product targets can also be used as target nucleic acids and are usually smaller than the original target.

- 30 The term "position" refers to a defined site on the nucleic acid. Such a position may, for example, be occupied by a nucleotide.

The term "allele" refers to any of the forms of the same gene that occur at the same locus on a homologous chromosome but differ in base sequence.

35



Nucleotides are referred to as being complementary if regular Watson-Crick base-pairing is ensured (e.g. G-C, A-T or A-U). This complementary base-pairing leads to matches whereas non-complementary base-pairing leads to mismatches. Any nucleobase creating such base-pairing (e.g. 7-deaza-dGTP, dITP, LNA-base analogues etc.) is also referred to  
 5 as being complementary.

For the realization of the method of the invention, the nucleic acids to be detected must be present in a form that is suitable for *in vitro* reaction. The sample to be examined, e.g. tissue sample, tissue extract, individual cells or a cell-containing fluid, is digested in a known  
 10 manner in order to break down the cell walls (e.g. thermal, enzymatic or chemical lysis or combinations thereof) to release the nucleic acids.

The sample is then brought into contact with diagnostic oligonucleotides of the invention. The conditions selected are such that the oligonucleotides of the invention hybridize with  
 15 the corresponding areas of the nucleic acids to be detected. This hybridization is generally known as annealing. Hybridization with irrelevant nucleic acids, i.e. those not to be detected, is avoided by selecting a suitable nucleotide sequence, length of the nucleotide, hybridisation temperature, adjuvants etc. (see Ausubel et. al in Current Protocols in Molecular Biology, pub. John Wiley & Sons (1998) and the details in the examples).

20 The number of different diagnostic oligonucleotides of the invention corresponds to at least the number of different target nucleic acids to be detected. The detection of two different mutations hence requires at least two different diagnostic oligonucleotides, each containing a nucleotide which is complementary to only one of the nucleotides which may  
 25 be present at the A position of the target nucleic acid. The diagnostic oligonucleotides used have a preferred length of less than 50 nucleotides, preferably in the range from 5 to 40 nucleotides, more preferably in the range from 5 to 30 nucleotides, even more preferably in the range from 5 to 25 nucleotides, e.g. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25. Each of the diagnostic oligonucleotides has a nucleotide se-  
 30 quence which is so highly complementary with respect to the area contiguous to position A that this oligonucleotide can hybridize with the nucleic acid to be detected.

The methods of the invention are improvements over the processes that are based on the principle of competitive priming or of the mismatch-priming-method and related methods.  
 35 However, regardless of the variant method applied, each diagnostic oligonucleotide used

in the detection of a target nucleic acid comprises at least one LNA. This LNA preferably corresponds to and is complementary to the nucleotide found at position A of the target nucleic acid to be detected.

5 One embodiment of the invention is a method for detecting in a sample the presence of a target nucleic acid Q' whose nucleotide sequence differs from a nucleic acid Q in at least one position A, the method comprising the steps of

- 10 a) combining nucleic acids present in the sample with an appropriate amount of nucleoside triphosphates, an agent for polymerisation of the nucleoside triphosphates and at least one diagnostic oligonucleotide containing at position A a nucleotide which is capable of hybridizing to the nucleotide at position A of nucleic acid Q' but not to a nucleotide at position A of nucleic acid Q under hybridisation conditions, the at least one diagnostic oligonucleotide containing at least one LNA,
- 15 b) extending any oligonucleotides which hybridize to the nucleic acids present in the sample to form extension products, wherein said nucleic acids are used as templates,
- 20 c) detecting any nucleic acids formed in step b) and thereby the presence of nucleic acid Q' in the sample.

This method can e.g. be applied to the reverse transcription of RNA into cDNA and other first strand synthesis applications. The method can further be made more efficient by cy-

25 cling the extension reaction by the addition of following steps instead of step c)

- c<sub>1</sub>) after the formation of extension products treating the reaction mixture under denaturing conditions to separate the extension products from the template,
- 30 d<sub>1</sub>) hybridising in the presence of an appropriate amount of nucleoside triphosphates and an agent for polymerisation of the nucleoside triphosphates the single stranded nucleic acids of step c<sub>1</sub>) with at least one diagnostic oligonucleotide containing at position A a nucleotide which is capable of hybridizing to the nucleotide at position A of nucleic acid Q' but not to a nucleotide at position A of nucleic acid

Q under hybridisation conditions, the at least one diagnostic oligonucleotide containing at least one LNA,

5 e<sub>1</sub>) repeating steps c<sub>1</sub>) through d<sub>1</sub>) a sufficient number of times to result in a detectable amount of extension products,

f<sub>1</sub>) detecting the extension products formed.

10 This method will result in a linear amplification and can further be extended to include re-amplification steps including a "downstream oligonucleotide" acting as primer for the synthesis of extension products of the opposite strand in the target nucleic acid which allows for an PCR-type of exponential amplification. The method for detecting the presence of a target nucleic acid Q' whose nucleotide sequence differs from a nucleic acid Q in at least one position A in a sample, will then comprise the steps of

- 15 a) combining the nucleic acids with an appropriate amount of nucleoside triphosphates, an agent for polymerisation of the nucleoside triphosphates, at least one downstream oligonucleotide and at least one diagnostic oligonucleotide containing at position A a nucleotide which is capable of hybridizing to the nucleotide at position A of nucleic acid Q' but not to a nucleotide at position A of nucleic acid Q under hybridisation conditions, the at least one diagnostic oligonucleotide containing at least one LNA,
- 20 b) extending any oligonucleotides which hybridize to the nucleic acids to form extension products, wherein said nucleic acids are used as templates,
- 25 c) after the formation of extension products treating the reaction mixture under denaturing conditions to separate the extension products from the template,
- 30 d) hybridising in the presence of an appropriate amount of nucleoside triphosphates and an agent for polymerisation of the nucleoside triphosphates the single stranded nucleic acids from step c) with at least one downstream oligonucleotide and at least one diagnostic oligonucleotide containing at position A a nucleotide which is capable of hybridizing to the nucleotide at position A of nucleic acid Q' but

not to a nucleotide at position A of nucleic acid Q under hybridisation conditions, the at least one diagnostic oligonucleotide containing at least one LNA,

5 e) repeating steps c) through d) a sufficient number of times to result in a detectable amount of extension products,

f) detecting the extension products formed.

Depending on which variant of the method in accordance with the invention is selected,  
10 the LNA-nucleotide(s) is/are located at a given site of the diagnostic oligonucleotides, preferably at position A, for example at internal positions of the diagnostic oligonucleotide or, preferably, at the 3'-end of the diagnostic oligonucleotides.

The presently described invention is particularly suited for multiplex analyses. In a number  
15 of situations genes are analysed which may contain a number of different mutations at different sites e.g. the TP53 gene having mutations in numerous codons. By including a number of carefully matched LNA-oligo/DNA-oligo pairs, each of which is directed towards a particular mutation and each of which may result in a characteristically sized fragment or characteristically labelled fragment, it is possible to discriminate between a number of  
20 different mutations by size-fractionating the fragments produced or detecting the labels of the fragments. The method for detecting target nucleic acids whose nucleotide sequences differ from one another in at least one position A in this case is similar to the above mentioned but here the diagnostic oligonucleotides consist of more than one sequence, the method thus comprises the steps of

25

a) combining the nucleic acids with an appropriate amount of nucleoside triphosphates, an agent for polymerisation of the nucleoside triphosphates and at least one set of diagnostic oligonucleotides under hybridisation conditions, the at least one set of diagnostic oligonucleotides having nucleotide sequences which differ  
30 from one another in at least one position A and contain at least one LNA,

b) extending any oligonucleotides which hybridize to the nucleic acids to form extension products, wherein said nucleic acids are used as templates,

35

c) detecting the nucleic acids formed in step b).

This constitutes a method which e.g. can be used as a sequence specific first strand synthesis of different nucleic acid sequences and can e.g. be used as a sequence specific reverse transcription of RNA. As also described above the method for detecting target nucleic acids whose nucleotide sequences differ from one another in at least one position A can be rendered more efficient by repeating the extension reaction a number of times as can be described by the addition of following steps instead of step c)

c<sub>1</sub>) after the formation of extension products treating the reaction mixture under denaturing conditions to separate the extension products from the template,

d<sub>1</sub>) hybridising in the presence of an appropriate amount of nucleoside triphosphates and an agent for polymerisation of the nucleoside triphosphates the single stranded nucleic acids from step c<sub>1</sub>) with at least one set of diagnostic oligonucleotides whose nucleotide sequences differ from one another in at least one position A to synthesise further extension products,

e<sub>1</sub>) repeating steps c<sub>1</sub>) through d<sub>1</sub>) a sufficient number of times to result in a detectable amount of extension products,

f<sub>1</sub>) detecting the extension products formed.

This method will result in a linear amplification of the different nucleic acids corresponding to the used oligonucleotides and can further be extended to include re-amplification steps including one or more downstream oligonucleotides acting as primers for the synthesis of the opposite strand in the target nucleic acid which allows for an PCR-type of exponential amplification. The method for detecting target nucleic acids whose nucleotide sequences differ from one another in at least one position A, will then comprise the steps of

a) combining the nucleic acids with an appropriate amount of nucleoside triphosphates, an agent for polymerisation of the nucleoside triphosphates, at least one downstream oligonucleotide and at least one set of diagnostic oligonucleotides under hybridisation conditions, the at least one set of diagnostic oligonucleotides having nucleotide sequences which differ from one another in at least one position A and contain at least one LNA,

b) extending any oligonucleotides which hybridize to the nucleic acids to form extension products, wherein said nucleic acids are used as templates,

5 c) after the formation of extension products treating the reaction mixture under denaturing conditions to separate the extension products from the template,

d) hybridising in the presence of an appropriate amount of nucleoside triphosphates and an agent for polymerisation of the nucleoside triphosphates the single stranded nucleic acids step c) with at least one downstream oligonucleotide and at  
10 at least one set of diagnostic oligonucleotides being oligonucleotides whose nucleotide sequences differ from one another in at least one position A to synthesise further extension products,

15 e) repeating steps c) through d) a sufficient number of times to result in a detectable amount of extension products,

f) detecting the extension products formed.

20 Also in these variants of the method it is preferred that the preferably 3'-terminal LNA-nucleotide of the diagnostic oligonucleotides corresponds to the position of the mutation (position A of the target nucleic acid) since the hybridization has a higher selectivity and specificity under these conditions.

25 It is preferred that the agent for polymerization in all the above mentioned methods is an enzyme e.g. a RNA polymerase, a revers transcriptase or a DNA polymerase. The DNA polymerase is preferably a thermostable DNA polymerase, e.g. Taq, Pfu, Pwo or Tth.

The principle of the detection of variant nucleic acids can be illustrated with the following  
30 example. One single reaction vessel contains a reagent mixture for a PCR for the simultaneous detection of two alleles of one gene. These alleles differ in one base position (position A of allele Q' (mutant) and position A of allele Q (normal)). Three primers, one generic primer complementary to a sequence found in both alleles and two primers each selective for only one of the two alleles, are used in the reaction. Primer selectivity is ef-  
35 fected in that the base at the 3'-end of the one selective primer is complementary to posi-

tion A of Q', the base of the other primer is complementary to position A of Q. Both primers contain a LNA at the 3'-position corresponding to position A. Moreover, in order to distinguish easily between the resulting PCR products, it is possible, for example to select the allele-specific primers such that PCR products of different lengths or with different labels are produced and conveniently detected by e.g. gel electrophoresis, capillary electrophoresis or various hybridizations techniques known in the art.

The sensitivity of the hybridization towards detection of differences in nucleotide sequence at a single-base may be enhanced through use of a ligase reaction. The observation in general (Nature 327, 293-297 (1987) and Nature 327, 298-303 (1987)) is that the shorter the length of a given oligonucleotide the better the discrimination towards single-base differences. However, short DNA-oligos have very low  $T_m$ , thus setting a lower limit to the length of DNA-oligos. On the other hand, the LNA-modification results in a very significant increase in the  $T_m$  (3-8°C pr. modification), therefore even short primers of 4-25 nucleotides are able to hybridise to a target nucleic acid at conditions compatible with standard PCR and LCR (Ligase Chain Reaction) conditions and furthermore do so with an improved discrimination towards single base differences. LCR, like PCR, uses multiple cycles of alternating temperature to amplify the numbers of a targeted sequence of DNA. LCR, however, does not use individual nucleotides for template extension but instead relies upon an excess of oligonucleotides which are complementary to both strands of the target region. Following the denaturation of a double stranded template DNA, the LCR procedure begins with the ligation of two (or more) oligonucleotide primers complementary to adjacent regions on one of the target strands. Oligonucleotides complementary to either strand can be joined. After ligation and a second denaturation step, the original template strands and the two (or more) newly joined products serve as templates for additional ligation to provide an exponential amplification of the targeted sequences. This method has been detailed in Genomics, 4:560-569 (1989), which is incorporated herein by reference. In the present invention it is proposed to include one short oligonucleotide containing at least one LNA preferably containing between 4 and 25 nucleotides, more preferably between 5 and 20 nucleotides, even more preferably between 6 and 15 nucleotides, e.g. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 nucleotides. This oligonucleotide is directed against gene-polymorphisms in the "LCR"-procedure. After multiple cycles of alternating temperature the numbers of a targeted sequence has been amplified and can be detected by gel electrophoresis, "sandwich"-hybridisation, or any of the many methods mentioned in the art. This is described by the following steps

a) combining the target nucleic acids with an appropriate amount of nucleoside triphosphates, at least two oligonucleotides wherein at least one of said oligonucleotides is a diagnostic oligonucleotide and an agent for ligation of the oligonucleotides under hybridisation conditions, the at least one diagnostic oligonucleotide being an oligonucleotide containing at position A a nucleotide which is complementary to the nucleotide found at position A of the target nucleic acid to be detected, said diagnostic oligonucleotide contains at least one LNA,

b) ligating any oligonucleotides which hybridize to the nucleic acids at adjacent positions to form ligation products, wherein said nucleic acids are used as templates,

c) treating the reaction mixture under denaturing conditions to separate the ligation products from the template after the ligation,

d) hybridising in the presence of an appropriate amount of nucleoside triphosphates and an agent for ligation of the oligonucleotides under hybridisation conditions the single stranded nucleic acids from step c) with at least one oligonucleotide being complementary to the ligation product from step b) and at least one diagnostic oligonucleotide being an oligonucleotide containing at position A a nucleotide which is complementary to the nucleotide found at position A of the target nucleic acid to be detected, said diagnostic oligonucleotide contains at least one LNA,

e) repeating steps c) through d) a sufficient number of times to result in a detectable amount of ligation products,

f) detecting the ligation products formed.

The method can of course be stopped after step b) and visualized immediately.

The preferred agent for ligation is an enzyme, e.g. a ligase, preferably a DNA-ligase, e.g. T4-DNA ligase, more preferably a thermostable ligase, e.g. Taq DNA ligase.



Another improvement of existing methods relates to the combined use of LCR and PCR. Accordingly it is proposed to hybridize two (or more) short (e.g. 8 bp) upstream LNA-primers to the template instead of one upstream primer. When two (or more) short LNA-primers are hybridizing next to each other they can be joined into one by a ligation reaction.

- 5 Thus the ligation will ensure sufficient specificity. Following the ligation the ligated primers can be extended. Since both thermostable ligase and polymerases have been described it is suggested to use multiple cycles of alternating temperature to amplify the numbers of a targeted sequence of DNA. Following the denaturation of a double stranded template DNA, the procedure begins with the ligation of two (or more) oligonucleotide primers complementary to adjacent regions on one of the target strands. The joined/ligated oligonucleotides form one upstream "primer" which is extended by the polymerase. After a second denaturation step, the original template strands and the newly formed products serve as templates for additional ligation and extension to provide an exponential amplification of the targeted sequences. After multiple cycles of alternating temperature the
- 10 15 numbers of a targeted sequence has been amplified.

The diagnostic oligonucleotides of all the above mentioned methods have the general formula of

20 a) 
$$5'-\text{Nu}_a(\text{Nu}_b\text{LNA}_c)_m\text{Nu}_d\text{A}_e\text{Nu}_f-3'$$

- where A is a LNA in position A in which the variant nucleic acid sequences of the target nucleic acids differs from one another; LNA is a LNA; Nu is a monomer selected from the group consisting of any nucleotides other than LNA capable of forming specific base-pairs with the variant nucleic acids; a, b, c, d and f are integers between 0 and 30, m is an integer between 1 and 8 and e is an integer between 1 and 6 with the proviso that the sum of
- 25 30 a, b, c, d, e and f is at least 5.

- This means that it is preferred that at least one position A of the diagnostic oligonucleotide sequence is complementary to position A of the target nucleic acid sequence to be detected but not to position A of the other target nucleic acid sequences. Furthermore, it is preferred that at least one position A in the diagnostic oligonucleotide is a LNA.
- 30

- The specific detection of the extension products, which is a measure for the presence or the quantity of the nucleic acid to be detected in the sample is, for example, also possible
- 35

by making use of the fact that the oligonucleotides used are discriminated by one additional feature. Such a distinction could be, for example, the varying lengths of oligonucleotides of one set. The extension of the different diagnostic oligonucleotides then produces extension products of different length. The diagnostic oligonucleotides can also be distinguished by having different labels. Such labels are, for example, colour or fluorescence molecules or chemical groups which can be detected in a subsequent reaction with the aid of a detectable group. Chemical groups of this kind include, e.g. haptens such as digoxin and digoxigenin. Haptens can be detected by reacting them with a labelled antibody to the hapten, the label is then detected. Another hapten could be biotin, for example, which can be detected by using a differently labelled antibody or streptavidin.

In the present context, the term "label" means a group which is detectable either by itself or as a part of an detection series. Examples of functional parts of reporter groups are biotin, digoxigenin, fluorescent groups (groups which are able to absorb electromagnetic radiation, e.g. light or X-rays, of a certain wavelength, and which subsequently reemits the energy absorbed as radiation of longer wavelength; illustrative examples are dansyl (5-dimethylamino)-1-naphthalenesulfonyl), DOXYL (N-oxy-4,4-dimethyloxazolidine), PROXYL (N-oxy-2,2,5,5-tetramethylpyrrolidine), TEMPO (N-oxy-2,2,6,6-tetramethylpiperidine), dinitrophenyl, acridines, coumarins, Cy3 and Cy5 (trademarks for Biological Detection Systems, Inc.), erythrosine, coumaric acid, umbelliferone, texas red, rhodamine, tetramethyl rhodamine, Rox, 7-nitrobenzo-2-oxa-1-diazole (NBD), pyrene, fluorescein, Europium, Ruthenium, Samarium, and other rare earth metals), radioisotopic labels, chemiluminescence labels (labels that are detectable via the emission of light during a chemical reaction), spin labels (a free radical (e.g. substituted organic nitroxides) or other paramagnetic probes (e.g.  $\text{Cu}^{2+}$ ,  $\text{Mg}^{2+}$ ) bound to a biological molecule being detectable by the use of electron spin resonance spectroscopy), enzymes (such as peroxidases, alkaline phosphatases,  $\beta$ -galactosidases, and glucose oxidases), antigens, antibodies, haptens (groups which are able to combine with an antibody, but which cannot initiate an immune response by itself, such as peptides and steroid hormones), carrier systems for cell membrane penetration such as: fatty acid residues, steroid moieties (cholesteryl), vitamin A, vitamin D, vitamin E, folic acid peptides for specific receptors, groups for mediating endocytosis, epidermal growth factor (EGF), bradykinin, and platelet derived growth factor (PDGF). Especially interesting examples are biotin, fluorescein, Texas Red, rhodamine, dinitrophenyl, digoxigenin, Ruthenium, Europium, Cy5, Cy3, etc.

The various extension products which are a measure for the quantity and the presence of the nucleic acid to be detected can be detected in various ways. They can be separately detected after separating the reagent mixture or sequentially once the reagent mixture is obtained or, provided appropriate labels are used, simultaneously according to known  
 5 methods. Preferably the individual oligonucleotides in each oligonucleotide set is labelled with detectable groups, said detectable groups being different for each individual diagnostic oligonucleotide.

The method described in EP-A-0 324 474 has proved to be the preferred method for the  
 10 use of steroid hormones as markers. Another method is the use of different fluorescent dyes. The oligonucleotides can either be labelled directly or antibodies to chemical groups can, for example, be provided with corresponding fluorescent labels. The various allelic products can be detected by simultaneously measuring the fluorescence in several channels corresponding to the number of different oligonucleotide primers. Also, part of the  
 15 products can be labelled with fluorescent labels whereas enzymes can be used for other products from the same reagent mixture. In principle it is possible to employ any known method of labelling and detecting.

In a sequential or simultaneous detection, the two different chemical groups, preferably  
 20 haptens which are detected by different antibodies, can be used, for example, to label the two oligonucleotides. Such haptens include digoxigenin, biotin, and fluorescein. Preferably, the antibody to fluorescein and the antibody to digoxigenin then have different enzymes as labels. The nucleic acids are then detected by sequential or simultaneous contact with enzyme-specific detectable substrates.

25

The mixture of the common extension reaction could also be subject to a denaturing reaction consequently using a capture probe to bind the single-stranded extension products to a solid phase which is immobilized or can be immobilized via a group (e.g. biotin). Alternatively, the capture probe could be permanently bound to a solid phase in advance e.g.  
 30 by anthraquinone photochemistry (WO 96/31557). In these procedures, the extension products of the reaction are immobilized together with all nucleic acids to be detected. Due to the specificity of the capture probe the nucleic acids can be detected sequentially (separated by a washing step). If the signal to be detected can be produced and detected at the same time in one vessel, such a washing step can be omitted.

35

In a variant two (or correspondingly more if detection involves more than two alleles and if more than two sets of specific primers are used) sufficient aliquots of liquid are taken from the reagent mixture after the extension reaction has occurred, and each aliquot is tested for extension products. The aliquots are taken from a multiplex reaction, and depending  
 5 on the number of different labels used for the oligonucleotides, a corresponding number of amplification products can be detected in each aliquot.

The method of the invention can be employed in a variety of applications. It is possible, for example, that the target nucleic acids to be detected from one single test subject (individual  
 10 ual diagnostics) are related to certain diseases, e.g. metabolic disorders, life style diseases, cancer or inheritable diseases. Examples of life style diseases are obesity, familial hypercholesterolaemia, atherosclerosis and diabetes.

The sample containing the nucleic acid to be detected will typically be cells, a tissue sam-  
 15 ple or a tissue extract of either archae, prokaryotic or eukaryotic origin. In the case the sample originates from an animal, e.g. a mammal or a human, the sample can be blood, serum, plasma, reticulocytes, lymphocytes, urine, bone marrow tissue, cerebrospinal fluid or any product prepared from blood or lymph or any type of tissue biopsy e.g. a muscle biopsy, a liver biopsy, a kidney biopsy, a bladder biopsy, a bone biopsy, a cartilage bi-  
 20 opsy, a skin biopsy, a pancreas biopsy, a biopsy of the intestinal tract, a thymus biopsy, a mammae biopsy, an uterus biopsy, a testicular biopsy, an eye biopsy or a brain biopsy.

Moreover, because of its good specificity, the method is also suited for assays where probes are subject to pool-screening. According to this principle a great number of indi-  
 25 vidual probes of different test subjects are mixed. When diagnosing the defect in the apo B 3500-gene, e.g. in connection with diabetes, familial hypercholesterolaemia, atherosclerosis or obesity, it has proven to be suitable to combine 64 probes in one pool. With the method of the invention it is possible to determine the presence or absence of this defect in a probe contained in the pool. When several partial pools from a pool with posi-  
 30 tive results are successively tested several times in accordance with the method of the invention, a small (as compared to prior art) number of determinations suffices to detect the gene which bears the defect.

In the method of the invention it has proven to be particularly advantageous to simultane-  
 35 ously detect the gene that does not bear the defect provided this gene is present. When

screening a pool, it is possible to alternatively detect the mutant product or, if no mutant allele is present, to detect the normal product. In addition to its function as a control for the reaction as such, the detection of both alleles is also a means of quantifying the nucleic acids to be detected.

5

Since the samples used in pool-screening methods are not patient-specific but anonymous, these methods can be used in the testing of diseases which are of particular epidemiological relevance, e.g. the contamination rate of the AIDS virus or, e.g. the frequency of elevated cholesterol levels, hypertension or diabetes (mass screening). These

10 methods serve to determine the frequency of alleles.

The same applies to the detection or diagnostics of particular species, sub-species or strains of organisms, micro-organisms or particular species, sub-species or strains of infectious agents especially of bacteria and in some cases to virus. The area of particular  
 15 interest focus on the simultaneous detection of closely related pathogenic/non-pathogenic species or strains. Further, the method of the invention is also suited to reliably detect e.g. E. coli K12 (evaluated in accordance with law of gene technology) via the known AT-sequences whereby various strands of E. coli are distinguished. The method can hence also be applied in analysis of environmental conditions.

20

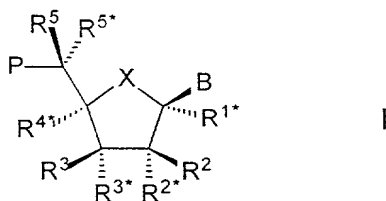
An important subject matter of the invention is the use of diagnostic oligonucleotides containing at least one LNA covalently attached to a solid surface e.g. by anthraquinone photochemistry (WO 96/31557) in e.g. an array format (Nature Genetics, suppl. vol. 21, Jan 1999, 1-60). Different diagnostic oligonucleotides containing at least one LNA di-  
 25 rected against different target nucleic acid sequences can be spotted in the array and permanently affixed to the solid surface. Such an array can subsequently be subjected to an extension reaction or to PCR. If the diagnostic oligonucleotides are attached to a suitable microscopically slide, e.g. surface-treated glass or polycarbonate slide, the extension reaction or the PCR reaction can be performed in one of the commercially available in situ  
 30 PCR thermocyclers. After PCR the positive spots (indicative of a specific mutation) can be detected by a suitable label. The result of such a procedure would be a semi-quantitative assessment of a large number of different target nucleic acids.

The main advantage of the invention over prior art is the possibility of simultaneously de-  
 35 tecting allelic compounds in mixtures contained in one single vessel with practically no

relevant secondary reactions. Digestions with restriction enzymes or subsequent hybridization is not required. This object is achieved by virtue of the fact that LNA oligomers obey the Watson-Crick base-pairing rules and form duplexes that are significantly more stable than similar duplexes formed by DNA:DNA, and that primers containing 3' located LNA(s) are substrate for the PCR reaction. The amplification processes of the individual alleles occur highly discriminative (cross reactions are not visible) and several reactions may take place in the same vessel.

Another subject matter of the invention is a reagent kit for the detection of nucleic acids. Said kit comprises a set of oligonucleotides in accordance with the invention and, if necessary, additional oligonucleotides and adjuvants necessary for the extension of the oligonucleotides. Further, the reagent kit can comprise additional oligonucleotides, for example, complementary strand primers. Preferably, the kit provides the enzyme and the oligonucleotides and the other reagents in separate containers.

When used herein, the term "LNA " or "LNA-containing oligonucleotide" refer to oligomers comprising at least one nucleoside analogue of the general formula I



wherein X is selected from -O-, -S-, -N(R<sup>N\*</sup>)-, -CR<sup>6</sup>(R<sup>6\*</sup>)-;

B is selected from nucleobases;

P designates the radical position for an internucleoside linkage to a succeeding monomer, or a 5'-terminal group, such internucleoside linkage or 5'-terminal group optionally including the substituent R<sup>5</sup>;

R<sup>3</sup> or R<sup>3\*</sup> is P\* which designates an internucleoside linkage to a preceding monomer, or a 3'-terminal group;

R<sup>4\*</sup> and R<sup>2\*</sup> together designates a biradical consisting of 1-4 groups/atoms selected from -C(R<sup>a</sup>R<sup>b</sup>)-, -C(R<sup>a</sup>)=C(R<sup>a</sup>)-, -C(R<sup>a</sup>)=N-, -O-, -Si(R<sup>a</sup>)<sub>2</sub>-, -S-, -SO<sub>2</sub>-, -N(R<sup>a</sup>)-, and >C=Z,

wherein Z is selected from -O-, -S-, and -N(R<sup>a</sup>)-, and R<sup>a</sup> and R<sup>b</sup> each is independently selected from hydrogen, optionally substituted C<sub>1-12</sub>-alkyl, optionally substituted C<sub>2-12</sub>-alkenyl, optionally substituted C<sub>2-12</sub>-alkynyl, hydroxy, C<sub>1-12</sub>-alkoxy, C<sub>2-12</sub>-alkenyloxy, carboxy, C<sub>1-12</sub>-alkoxycarbonyl, C<sub>1-12</sub>-alkylcarbonyl, formyl, aryl, aryloxy-carbonyl, aryloxy, arylcarbonyl, heteroaryl, heteroaryloxy-carbonyl, heteroaryloxy, heteroarylcarbonyl, amino, mono- and di(C<sub>1-6</sub>-alkyl)amino, carbamoyl, mono- and di(C<sub>1-6</sub>-alkyl)-amino-carbonyl, amino-C<sub>1-6</sub>-alkyl-aminocarbonyl, mono- and di(C<sub>1-6</sub>-alkyl)amino-C<sub>1-6</sub>-alkyl-aminocarbonyl, C<sub>1-6</sub>-alkyl-carbonylamino, carbamido, C<sub>1-6</sub>-alkanoyloxy, sulphono, C<sub>1-6</sub>-alkylsulphonyloxy, nitro, azido, sulphanyl, C<sub>1-6</sub>-alkylthio, halogen, DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands, where aryl and heteroaryl may be optionally substituted, and where two geminal substituents R<sup>a</sup> and R<sup>b</sup> together may designate optionally substituted methylene (=CH<sub>2</sub>, optionally substituted one or two times with substituents as defined as optional substituents for aryl); and

each of the substituents R<sup>1\*</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>3\*</sup>, R<sup>5</sup>, R<sup>5\*</sup>, R<sup>6</sup> and R<sup>6\*</sup> which are present and not involved in P or P\*, is independently selected from hydrogen, optionally substituted C<sub>1-12</sub>-alkyl, optionally substituted C<sub>2-12</sub>-alkenyl, optionally substituted C<sub>2-12</sub>-alkynyl, hydroxy, C<sub>1-12</sub>-alkoxy, C<sub>2-12</sub>-alkenyloxy, carboxy, C<sub>1-12</sub>-alkoxycarbonyl, C<sub>1-12</sub>-alkylcarbonyl, formyl, aryl, aryloxy-carbonyl, aryloxy, arylcarbonyl, heteroaryl, heteroaryloxy-carbonyl, heteroaryloxy, heteroarylcarbonyl, amino, mono- and di(C<sub>1-6</sub>-alkyl)amino, carbamoyl, mono- and di(C<sub>1-6</sub>-alkyl)-amino-carbonyl, amino-C<sub>1-6</sub>-alkyl-aminocarbonyl, mono- and di(C<sub>1-6</sub>-alkyl)amino-C<sub>1-6</sub>-alkyl-aminocarbonyl, C<sub>1-6</sub>-alkyl-carbonylamino, carbamido, C<sub>1-6</sub>-alkanoyloxy, sulphono, C<sub>1-6</sub>-alkylsulphonyloxy, nitro, azido, sulphanyl, C<sub>1-6</sub>-alkylthio, halogen, DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands (where the latter groups may include a spacer as defined for the substituent B), where aryl and heteroaryl may be optionally substituted, and where two geminal substituents together may designate oxo, thioxo, imino, or optionally substituted methylene, or together may form a spiro biradical consisting of a 1-5 carbon atom(s) alkylene chain which is optionally interrupted and/or terminated by one or more heteroatoms/groups selected from -O-, -S-, and -(NR<sup>N</sup>)- where R<sup>N</sup> is selected from hydrogen and C<sub>1-4</sub>-alkyl, and where two adjacent (non-geminal) substituents may designate an additional bond resulting in a double bond; and R<sup>N\*</sup>, when present and not involved in a biradical, is selected from hydrogen and C<sub>1-4</sub>-alkyl;

and basic salts and acid addition salts thereof.

When used herein, the term "LNA" (Locked Nucleoside Analogues) refers to the bi-cyclic  
 5 nucleoside analogues incorporated in the oligomer (general formula I).

In the present context, the terms "nucleobase" covers naturally occurring nucleobases as well as non-naturally occurring nucleobases. It should be clear to the person skilled in the art that various nucleobases which previously have been considered "non-naturally occurring"  
 10 ring" have subsequently been found in nature. Thus, "nucleobase" includes not only the known purine and pyrimidine heterocycles, but also heterocyclic analogues and tautomers thereof. Illustrative examples of nucleobases are adenine, guanine, thymine, cytosine, uracil, purine, xanthine, diaminopurine, 8-oxo-N<sup>6</sup>-methyladenine, 7-deazaxanthine, 7-deazaguanine, N<sup>4</sup>,N<sup>4</sup>-ethanocytosin, N<sup>6</sup>,N<sup>6</sup>-ethano-2,6-diaminopurine, 5-methylcytosine,  
 15 5-(C<sup>3</sup>-C<sup>6</sup>)-alkynylcytosine, 5-fluorouracil, 5-bromouracil, pseudoisocytosine, 2-hydroxy-5-methyl-4-triazolopyridin, isocytosine, isoguanin, inosine and the "non-naturally occurring" nucleobases described in Benner et al., U.S. Pat No. 5,432,272. The term "nucleobase" is intended to cover every and all of these examples as well as analogues and tautomers thereof. Especially interesting nucleobases are adenine, guanine, thymine, cytosine, and  
 20 uracil, which are considered as the naturally occurring nucleobases in relation to therapeutic and diagnostic application in humans.

When used herein, the term "DNA intercalator" means a group which can intercalate into a DNA or RNA helix, duplex or triplex. Examples of functional parts of DNA intercalators  
 25 are acridines, anthracene, quinones such as anthraquinone, indole, quinoline, isoquinoline, dihydroquinones, anthracyclines, tetracyclines, methylene blue, anthracyclinone, psoralens, coumarins, ethidium-halides, dynemicin, metal complexes such as 1,10-phenanthroline-copper, tris(4,7-diphenyl-1,10-phenanthroline)ruthenium-cobalt-enediynes such as calcheamicin, porphyrins, distamycin, netropcin, viologen, daunomycin. Especially  
 30 interesting examples are acridines, quinones such as anthraquinone, methylene blue, psoralens, coumarins, and ethidium-halides.

In the present context, the term "photochemically active groups" covers compounds which are able to undergo chemical reactions upon irradiation with light. Illustrative examples of  
 35 functional groups hereof are quinones, especially 6-methyl-1,4-naphthoquinone, anthraqui-



none, naphthoquinone, and 1,4-dimethyl-anthraquinone, diazirines, aromatic azides, benzophenones, psoralens, diazo compounds, and diazirino compounds.

In the present context "thermochemically reactive group" is defined as a functional group which is able to undergo thermochemically-induced covalent bond formation with other groups. Illustrative examples of functional parts thermochemically reactive groups are carboxylic acids, carboxylic acid esters such as activated esters, carboxylic acid halides such as acid fluorides, acid chlorides, acid bromide, and acid iodides, carboxylic acid azides, carboxylic acid hydrazides, sulfonic acids, sulfonic acid esters, sulfonic acid halides, semicarbazides, thiosemicarbazides, aldehydes, ketones, primary alcohols, secondary alcohols, tertiary alcohols, phenols, alkyl halides, thiols, disulphides, primary amines, secondary amines, tertiary amines, hydrazines, epoxides, maleimides, and boronic acid derivatives.

In the present context, the term "chelating group" means a molecule that contains more than one binding site and frequently binds to another molecule, atom or ion through more than one binding site at the same time. Examples of functional parts of chelating groups are iminodiacetic acid, nitrilotriacetic acid, ethylenediamine tetraacetic acid (EDTA), aminophosphonic acid, etc.

In the present context "ligand" means something which binds. Ligands can comprise functional groups such as: aromatic groups (such as benzene, pyridine, naphthalene, anthracene, and phenanthrene), heteroaromatic groups (such as thiophene, furan, tetrahydrofuran, pyridine, dioxane, and pyrimidine), carboxylic acids, carboxylic acid esters, carboxylic acid halides, carboxylic acid azides, carboxylic acid hydrazides, sulfonic acids, sulfonic acid esters, sulfonic acid halides, semicarbazides, thiosemicarbazides, aldehydes, ketones, primary alcohols, secondary alcohols, tertiary alcohols, phenols, alkyl halides, thiols, disulphides, primary amines, secondary amines, tertiary amines, hydrazines, epoxides, maleimides, C<sub>1</sub>-C<sub>20</sub> alkyl groups optionally interrupted or terminated with one or more heteroatoms such as oxygen atoms, nitrogen atoms, and/or sulphur atoms, optionally containing aromatic or mono/polyunsaturated hydrocarbons, polyoxyethylene such as polyethylene glycol, oligo/polyamides such as poly-β-alanine, polyglycine, polylysine, peptides, oligo/polysaccharides, oligo/polyphosphates, toxins, antibiotics, cell poisons, and steroids, and also "affinity ligands", i.e. functional groups or biomolecules that have a specific affinity for sites on particular proteins, antibodies, poly- and oligosaccharides, and other biomolecules.

It will be clear for the person skilled in the art that the above-mentioned specific examples under DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands correspond to the "active/functional" part of the groups in question. For the person skilled in the art it is furthermore clear that DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands are typically represented in the form M-K- where M is the "active/functional" part of the group in question and where K is a spacer through which the "active/functional" part is attached to the 5- or 6-membered ring. Thus, it should be understood that the group B, in the case where B is selected from DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands, has the form M-K-, where M is the "active/functional" part of the DNA intercalator, photochemically active group, thermochemically active group, chelating group, reporter group, and ligand, respectively, and where K is an optional spacer comprising 1-50 atoms, preferably 1-30 atoms, in particular 1-15 atoms, between the 5- or 6-membered ring and the "active/functional" part.

In the present context, the term "spacer" means a thermochemically and photochemically non-active distance-making group and is used to join two or more different moieties of the types defined above. Spacers are selected on the basis of a variety of characteristics including their hydrophobicity, hydrophilicity, molecular flexibility and length (e.g. see Hermanson et. al., "Immobilized Affinity Ligand Techniques", Academic Press, San Diego, California (1992), p. 137-ff). Generally, the length of the spacers are less than or about 400 Å, in some applications preferably less than 100 Å. The spacer, thus, comprises a chain of carbon atoms optionally interrupted or terminated with one or more heteroatoms, such as oxygen atoms, nitrogen atoms, and/or sulphur atoms. Thus, the spacer K may comprise one or more amide, ester, amino, ether, and/or thioether functionalities, and optionally aromatic or mono/polyunsaturated hydrocarbons, polyoxyethylene such as polyethylene glycol, oligo/polyamides such as poly-β-alanine, polyglycine, polylysine, and peptides in general, oligosaccharides, oligo/polyphosphates. Moreover the spacer may consist of combined units thereof. The length of the spacer may vary, taking into consideration the desired or necessary positioning and spatial orientation of the "active/functional" part of the group in question in relation to the 5- or 6-membered ring. In particularly interesting embodiments, the spacer includes a chemically cleavable group.

Examples of such chemically cleavable groups include disulphide groups cleavable under reductive conditions, peptide fragments cleavable by peptidases, etc.

In one variant, K designates a single bond so that the "active/functional" part of the group  
5 in question is attached directly to the 5- or 6-membered ring.

In a preferred embodiment, the substituent B in the general formulae I and II is preferably selected from nucleobases, in particular from adenine, guanine, thymine, cytosine and urasil.

10

In the oligomers (formula I), P designates the radical position for an internucleoside linkage to a succeeding monomer, or a 5'-terminal group. The first possibility applies when the LNA in question is not the 5'-terminal "monomer", whereas the latter possibility applies when the LNA in question is the 5'-terminal "monomer". It should be understood (which  
15 also will be clear from the definition of internucleoside linkage and 5'-terminal group further below) that such an internucleoside linkage or 5'-terminal group may include the substituent  $R^5$  (or equally applicable: the substituent  $R^{5*}$ ) thereby forming a double bond to the group P. (5'-Terminal refers to the position corresponding to the 5' carbon atom of a ribose moiety in a nucleoside.)

20

On the other hand, an internucleoside linkage to a preceding monomer or a 3'-terminal group ( $P^*$ ) may originate from the positions defined by one of the substituents  $R^3$  or  $R^{3*}$ , preferably from the positions defined by the substituents  $R^{3*}$ . (3'-Terminal refers to the position corresponding to the 3' carbon atom of a ribose moiety in a nucleoside.)

25

It should be understood that the orientation of the group  $P^*$  either as  $R^{3*}$  ("normal" configuration) or as  $R^3$  (xylo configuration) represent two equally interesting possibilities. It has been found that all-"normal" ( $R^{3*} = P^*$ ) oligomers and oligomers with combinations of "normal" LNA monomers and nucleotides (2-deoxynucleotides and/or nucleotides) hybridise strongly (with increasing affinity) to DNA, RNA and other LNA oligomers. It is presently believed that combination of all-xylo LNA oligomers and oligomers with xylo LNA ( $R^3 = P^*$ ) monomers and, e.g., xylo nucleotides (nucleotides and/or 2-deoxynucleotides) will give rise to comparable hybridisation properties. It has been shown that an oligomer with "normal" configuration ( $R^{3*} = P^*$ ) will give rise to an anti-parallel orientation of an LNA  
30 oligomer when hybridised (with increasing affinity) to either DNA, RNA or another LNA  
35

oligomer. It is thus contemplated that an oligomer with xylo configuration ( $R^3 = P^*$ ) will give rise to a parallel orientation when hybridised to DNA, RNA or another LNA.

In view of the above, it is contemplated that the combination of "normal" LNAs and xylo-  
 5 LNAs in one oligomer can give rise to interesting properties as long as these monomers of different type are located in domains, i.e. so that an uninterrupted domain of at least 5, such as at least 10, monomers (e.g. xylo-LNA, xylo-nucleotides, etc. monomers) is followed by an uninterrupted domain of at least 5, e.g. at least 10, monomers of the other type (e.g. "normal" LNA, "normal" nucleotides, etc.), etc. Such chimeric type oligomers  
 10 may, e.g., be used to capture nucleic acids.

In the present context, the term "monomer" relates to naturally occurring nucleosides, non-naturally occurring nucleosides, PNAs, etc. as well as LNAs. Thus, the term "succeeding monomer" relates to the neighbouring monomer in the 5'-terminal direction and  
 15 the "preceding monomer" relates to the neighbouring monomer in the 3'-terminal direction. Such succeeding and preceding monomers, seen from the position of an LNA monomer, may be naturally occurring nucleosides or non-naturally occurring nucleosides, or even further LNA monomers.

20 Consequently, in the present context (as can be derived from the definitions above), the term "oligomer" means an oligonucleotide modified by the incorporation of one or more LNA(s).

In the present context, the orientation of the biradical ( $R^{2*}-R^{4*}$ ) is so that the left-hand side  
 25 represents the substituent with the lowest number and the right-hand side represents the substituent with the highest number, thus, when  $R^{2*}$  and  $R^{4*}$  together designate a biradical "-O-CH<sub>2</sub>-", it is understood that the oxygen atom represents  $R^{2*}$ , thus the oxygen atom is e.g. attached to the position of  $R^{2*}$ , and the methylene group represents  $R^{4*}$ .

30 Considering the numerous interesting possibilities for the structure of the biradical ( $R^{2*}-R^{4*}$ ) in LNA(s) incorporated in oligomers, it is believed that the biradical preferably is selected from  $-(CR^*R^*)_r-Y-(CR^*R^*)_s-$ ,  $-(CR^*R^*)_r-Y-(CR^*R^*)_s-Y-$ ,  $-Y-(CR^*R^*)_{r+s}-Y-$ ,  $-Y-(CR^*R^*)_r-Y-(CR^*R^*)_s-$ ,  $-(CR^*R^*)_{r+s}-Y-$ ,  $-Y-Y-$ , wherein each Y is independently selected from -O-, -S-, -Si( $R^*$ )<sub>2</sub>-, -N( $R^*$ )-, >C=O, -C(=O)-N( $R^*$ )-, and -N( $R^*$ )-C(=O)-, each  $R^*$  is independently se-  
 35 lected from hydrogen, halogen, azido, cyano, nitro, hydroxy, mercapto, amino, mono- or

- di(C<sub>1-6</sub>-alkyl)amino, optionally substituted C<sub>1-6</sub>-alkoxy, optionally substituted C<sub>1-6</sub>-alkyl, DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands, and/or two adjacent (non-geminal) R<sup>\*</sup> may together designate a double bond; and each of r and s is 0-4 with the proviso that the sum r+s is 1-4. Particularly interesting situations are those wherein the biradical is selected from -Y-, -(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-, -(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r</sub>-Y-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>s</sub>-, and -Y-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-Y-, wherein and each of r and s is 0-3 with the proviso that the sum r+s is 1-4.

- Particularly interesting oligomers are those wherein R<sup>2\*</sup> and R<sup>4\*</sup> in at least one LNA in the oligomer together designate a biradical selected from -O-, -S-, -N(R<sup>\*</sup>)-, -(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s+1</sub>-, -(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r</sub>-O-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>s</sub>-, -(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r</sub>-S-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>s</sub>-, -(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r</sub>-N(R<sup>\*</sup>)-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>s</sub>-, -O-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-O-, -S-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-O-, -O-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-S-, -N(R<sup>\*</sup>)-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-O-, -O-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-N(R<sup>\*</sup>)-, -S-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-S-, -N(R<sup>\*</sup>)-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-N(R<sup>\*</sup>)-, -N(R<sup>\*</sup>)-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-S-, and -S-(CR<sup>\*</sup>R<sup>\*</sup>)<sub>r+s</sub>-N(R<sup>\*</sup>)-

- It is furthermore preferred that one R<sup>\*</sup> is selected from hydrogen, hydroxy, optionally substituted C<sub>1-6</sub>-alkoxy, optionally substituted C<sub>1-6</sub>-alkyl, DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands, and any remaining substituents R<sup>\*</sup> are hydrogen.
- In one preferred variant, one group R<sup>\*</sup> in the biradical of at least one LNA is selected from DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands (where the latter groups may include a spacer as defined for the substituent B).
- Preferably, each of the substituents R<sup>1\*</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>3\*</sup>, R<sup>5</sup>, R<sup>6\*</sup>, R<sup>6</sup> and R<sup>6\*</sup> of the LNA(s), which are present and not involved in P or P<sup>\*</sup>, is independently selected from hydrogen, optionally substituted C<sub>1-6</sub>-alkyl, optionally substituted C<sub>2-6</sub>-alkenyl, hydroxy, C<sub>1-6</sub>-alkoxy, C<sub>2-6</sub>-alkenyloxy, carboxy, C<sub>1-6</sub>-alkoxycarbonyl, C<sub>1-6</sub>-alkylcarbonyl, formyl, amino, mono- and di(C<sub>1-6</sub>-alkyl)amino, carbamoyl, mono- and di(C<sub>1-6</sub>-alkyl)-amino-carbonyl, C<sub>1-6</sub>-alkyl-carbonylamino, carbamido, azido, C<sub>1-6</sub>-alkanoyloxy, sulphonyl, sulphanyl, C<sub>1-6</sub>-alkylthio, DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands, and halogen, where two geminal substituents together may designate oxo, and where R<sup>N\*</sup>, when present and not involved in a biradical, is selected from hydrogen and C<sub>1-4</sub>-alkyl.

35

In a preferred variant of the LNAs, X is selected from -O-, -S-, and -NR<sup>N\*</sup>-, in particular -O-, and each of the substituents R<sup>1\*</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>3\*</sup>, R<sup>5</sup>, R<sup>6</sup>, R<sup>6\*</sup> and R<sup>6\*</sup> of the LNA(s), which are present and not involved in P or P\*, designate hydrogen.

- 5 In an even more preferred variant, X is O, R<sup>2</sup> selected from hydrogen, hydroxy, and optionally substituted C<sub>1-6</sub>-alkoxy, one of R<sup>3</sup> and R<sup>3\*</sup> is P\* and the other is hydrogen, and R<sup>1\*</sup>, R<sup>5</sup>, and R<sup>5\*</sup> designate hydrogen, and, more specifically, the biradical (R<sup>2\*</sup>-R<sup>4\*</sup>) is selected from -O-, -(CH<sub>2</sub>)<sub>0-1</sub>-O-(CH<sub>2</sub>)<sub>1-3</sub>-, -(CH<sub>2</sub>)<sub>0-1</sub>-S-(CH<sub>2</sub>)<sub>1-3</sub>-, -(CH<sub>2</sub>)<sub>0-1</sub>-N(R<sup>N</sup>)-(CH<sub>2</sub>)<sub>1-3</sub>-, and -(CH<sub>2</sub>)<sub>2-4</sub>-, in particular from -O-CH<sub>2</sub>-, -S-CH<sub>2</sub>-, and -NR<sup>H</sup>-CH<sub>2</sub>-. Generally, with due regard to the
- 10 results obtained so far, it is preferred that the biradical constituting R<sup>2\*</sup> and R<sup>4\*</sup> forms a two carbon atom bridge, *i.e.* the biradical forms a five membered ring with the furanose ring (X=O). Particularly interesting are also those oligomers where R<sup>2\*</sup> and R<sup>4\*</sup> of an incorporated LNA of formula I together designate a biradical selected from -O-CH<sub>2</sub>-, -S-CH<sub>2</sub>-, and -NR<sup>H</sup>-CH<sub>2</sub>-; X is O, B designates a nucleobase selected from adenine, guanine, thymine,
- 15 cytosine and urasil; R<sup>2</sup> is hydrogen, one of R<sup>3</sup> or R<sup>3\*</sup> designates P\* and the other is hydrogen, and R<sup>1\*</sup>, R<sup>3</sup>, R<sup>5</sup>, and R<sup>5\*</sup> designate hydrogen.

- In these embodiments, it is furthermore preferred that at least one LNA incorporated in an oligomer includes a nucleobase (substituent B) selected from adenine and guanine. In
- 20 particular, it is preferred that an oligomer have LNA incorporated therein both include at least one nucleobase selected from thymine, urasil and cytosine and at least one nucleobase selected from adenine and guanine. For LNA monomers, it is especially preferred that the nucleobase is selected from adenine and guanine.

- 25 Within a variant of these interesting embodiments, all monomers of a oligonucleotide are LNA monomers.

As it will be evident from the general formula I (LNA(s) in an oligomer) and the definitions associated therewith, there may be one or several asymmetric carbon atoms present in

30 the oligomers depending on the nature of the substituents and possible biradicals, cf. below.

In one variant, R<sup>3\*</sup> designates P\*. In another variant, R<sup>3</sup> designates P\*, and in a third variant, some R<sup>3\*</sup> designates P\* in some LNAs and R<sup>3</sup> designates P\* in other LNAs within an

35 oligomer.

The oligomers typically comprise 1-10000 LNA(s) of the general formula I and 0-10000 nucleosides selected from naturally occurring nucleosides and nucleoside analogues. The sum of the number of nucleosides and the number of LNA(s) is at least 2, preferably at least 3, in particular at least 5, especially at least 7, such as in the range of 2-15000, preferably in the range of 2-100, such as 3-100, in particular in the range of 2-50, such as 3-50 or 5-50 or 7-50.

In the present context, the term "nucleoside" means a glycoside of a heterocyclic base.

- 10 The term "nucleoside" is used broadly as to include non-naturally occurring nucleosides, naturally occurring nucleosides as well as other nucleoside analogues. Illustrative examples of nucleosides are ribonucleosides comprising a ribose moiety as well as deoxyribonucleosides comprising a deoxyribose moiety. With respect to the bases of such nucleosides, it should be understood that this may be any of the naturally occurring bases, *e.g.*
- 15 adenine, guanine, cytosine, thymine, and uracil, as well as any modified variants thereof or any possible unnatural bases.

When considering the definitions and the known nucleosides (naturally occurring and non-naturally occurring) and nucleoside analogues (including known bi- and tricyclic ana-

- 20 logues), it is clear that an oligomer may comprise one or more LNA(s) (which may be identical or different both with respect to the selection of substituent and with respect to selection of biradical) and one or more nucleosides and/or nucleoside analogues. In the present context "oligonucleotide" means a successive chain of nucleosides connected via internucleoside linkages, however, it should be understood that a nucleobase in one or
- 25 more nucleotide units (monomers) in an oligomer (oligonucleotide) may have been modified with a substituent B as defined above.

As mentioned above, the LNA(s) of an oligomer are connected with other monomers via an internucleoside linkage. In the present context, the term "internucleoside linkage"

- 30 means a linkage consisting of 2 to 4, preferably 3, groups/atoms selected from  $-\text{CH}_2-$ ,  $-\text{O}-$ ,  $-\text{S}-$ ,  $-\text{NR}^{\text{H}}-$ ,  $>\text{C}=\text{O}$ ,  $>\text{C}=\text{NR}^{\text{H}}$ ,  $>\text{C}=\text{S}$ ,  $-\text{Si}(\text{R}'')_2-$ ,  $-\text{SO}-$ ,  $-\text{S}(\text{O})_2-$ ,  $-\text{P}(\text{O})_2-$ ,  $-\text{PO}(\text{BH}_3)-$ ,  $-\text{P}(\text{O},\text{S})-$ ,  $-\text{P}(\text{S})_2-$ ,  $-\text{PO}(\text{R}'')-$ ,  $-\text{PO}(\text{OCH}_3)-$ , and  $-\text{PO}(\text{NHR}^{\text{H}})-$ , where  $\text{R}^{\text{H}}$  is selected from hydrogen and  $\text{C}_{1-4}$ -alkyl, and  $\text{R}''$  is selected from  $\text{C}_{1-6}$ -alkyl and phenyl. Illustrative examples of such internucleoside linkages are  $-\text{CH}_2-\text{CH}_2-\text{CH}_2-$ ,  $-\text{CH}_2-\text{CO}-\text{CH}_2-$ ,  $-\text{CH}_2-\text{CHOH}-\text{CH}_2-$ ,  $-\text{O}-\text{CH}_2-\text{O}-$ ,
- 35  $-\text{O}-\text{CH}_2-\text{CH}_2-$ ,  $-\text{O}-\text{CH}_2-\text{CH}=($  (including  $\text{R}^5$  when used as a linkage to a succeeding mono-

mer),  $-\text{CH}_2\text{-CH}_2\text{-O-}$ ,  $-\text{NR}^{\text{H}}\text{-CH}_2\text{-CH}_2\text{-}$ ,  $-\text{CH}_2\text{-CH}_2\text{-NR}^{\text{H}}\text{-}$ ,  $-\text{CH}_2\text{-NR}^{\text{H}}\text{-CH}_2\text{-}$ ,  $-\text{O-CH}_2\text{-CH}_2\text{-NR}^{\text{H}}\text{-}$ ,  $-\text{NR}^{\text{H}}\text{-CO-O-}$ ,  $-\text{NR}^{\text{H}}\text{-CO-NR}^{\text{H}}\text{-}$ ,  $-\text{NR}^{\text{H}}\text{-CS-NR}^{\text{H}}\text{-}$ ,  $-\text{NR}^{\text{H}}\text{-C(=NR}^{\text{H}}\text{)-NR}^{\text{H}}\text{-}$ ,  $-\text{NR}^{\text{H}}\text{-CO-CH}_2\text{-NR}^{\text{H}}\text{-}$ ,  $-\text{O-CO-O-}$ ,  $-\text{O-CO-CH}_2\text{-O-}$ ,  $-\text{O-CH}_2\text{-CO-O-}$ ,  $-\text{CH}_2\text{-CO-NR}^{\text{H}}\text{-}$ ,  $-\text{O-CO-NR}^{\text{H}}\text{-}$ ,  $-\text{NR}^{\text{H}}\text{-CO-CH}_2\text{-}$ ,  $-\text{O-CH}_2\text{-CO-NR}^{\text{H}}\text{-}$ ,  $-\text{O-CH}_2\text{-CH}_2\text{-NR}^{\text{H}}\text{-}$ ,  $-\text{CH=N-O-}$ ,  $-\text{CH}_2\text{-NR}^{\text{H}}\text{-O-}$ ,  $-\text{CH}_2\text{-O-N=}$  (including  $\text{R}^5$  when used as a linkage to a succeeding monomer),  $-\text{CH}_2\text{-O-NR}^{\text{H}}\text{-}$ ,  $-\text{CO-NR}^{\text{H}}\text{-CH}_2\text{-}$ ,  $-\text{CH}_2\text{-NR}^{\text{H}}\text{-O-}$ ,  $-\text{CH}_2\text{-NR}^{\text{H}}\text{-CO-}$ ,  $-\text{O-NR}^{\text{H}}\text{-CH}_2\text{-}$ ,  $-\text{O-NR}^{\text{H}}\text{-}$ ,  $-\text{O-CH}_2\text{-S-}$ ,  $-\text{S-CH}_2\text{-O-}$ ,  $-\text{CH}_2\text{-CH}_2\text{-S-}$ ,  $-\text{O-CH}_2\text{-CH}_2\text{-S-}$ ,  $-\text{S-CH}_2\text{-CH=}$  (including  $\text{R}^5$  when used as a linkage to a succeeding monomer),  $-\text{S-CH}_2\text{-CH}_2\text{-}$ ,  $-\text{S-CH}_2\text{-CH}_2\text{-O-}$ ,  $-\text{S-CH}_2\text{-CH}_2\text{-S-}$ ,  $-\text{CH}_2\text{-S-CH}_2\text{-}$ ,  $-\text{CH}_2\text{-SO-CH}_2\text{-}$ ,  $-\text{CH}_2\text{-SO}_2\text{-CH}_2\text{-}$ ,  $-\text{O-SO-O-}$ ,  $-\text{O-S(O)}_2\text{-O-}$ ,  $-\text{O-S(O)}_2\text{-CH}_2\text{-}$ ,  $-\text{O-S(O)}_2\text{-NR}^{\text{H}}\text{-}$ ,  $-\text{NR}^{\text{H}}\text{-S(O)}_2\text{-CH}_2\text{-}$ ,  $-\text{O-S(O)}_2\text{-CH}_2\text{-}$ ,  $-\text{O-P(O)}_2\text{-O-}$ ,  $-\text{O-P(O,S)-O-}$ ,  $-\text{O-P(S)}_2\text{-O-}$ ,  $-\text{S-P(O)}_2\text{-O-}$ ,  $-\text{S-P(O,S)-O-}$ ,  $-\text{S-P(S)}_2\text{-O-}$ ,  $-\text{O-P(O)}_2\text{-S-}$ ,  $-\text{O-P(O,S)-S-}$ ,  $-\text{O-P(S)}_2\text{-S-}$ ,  $-\text{S-P(O)}_2\text{-S-}$ ,  $-\text{S-P(O,S)-S-}$ ,  $-\text{S-P(S)}_2\text{-S-}$ ,  $-\text{O-PO(R'')-O-}$ ,  $-\text{O-PO(OCH}_3\text{)-O-}$ ,  $-\text{O-PO(OCH}_2\text{CH}_3\text{)-O-}$ ,  $-\text{O-PO(OCH}_2\text{CH}_2\text{S-R)-O-}$ ,  $-\text{O-PO(BH}_3\text{)-O-}$ ,  $-\text{O-PO(NHR}^{\text{N}}\text{)-O-}$ ,  $-\text{O-P(O)}_2\text{-NR}^{\text{H}}\text{-}$ ,  $-\text{NR}^{\text{H}}\text{-P(O)}_2\text{-O-}$ ,  $-\text{O-P(O,NR}^{\text{H}}\text{)-O-}$ ,  $-\text{CH}_2\text{-P(O)}_2\text{-O-}$ ,  $-\text{O-P(O)}_2\text{-CH}_2\text{-}$ , and  $-\text{O-Si(R'')}_2\text{-O-}$ ; among which  $-\text{CH}_2\text{-CO-NR}^{\text{H}}\text{-}$ ,  $-\text{CH}_2\text{-NR}^{\text{H}}\text{-O-}$ ,  $-\text{S-CH}_2\text{-O-}$ ,  $-\text{O-P(O)}_2\text{-O-}$ ,  $-\text{O-P(O,S)-O-}$ ,  $-\text{O-P(S)}_2\text{-O-}$ ,  $-\text{NR}^{\text{H}}\text{-P(O)}_2\text{-O-}$ ,  $-\text{O-P(O,NR}^{\text{H}}\text{)-O-}$ ,  $-\text{O-PO(R'')-O-}$ ,  $-\text{O-PO(CH}_3\text{)-O-}$ , and  $-\text{O-PO(NHR}^{\text{N}}\text{)-O-}$ , where  $\text{R}^{\text{H}}$  is selected from hydrogen and  $\text{C}_{1-4}$ -alkyl, and  $\text{R}''$  is selected from  $\text{C}_{1-6}$ -alkyl and phenyl, are especially preferred. Further illustrative examples are given in Mesmaeker et. al., Current Opinion in Structural Biology 1995, 5, 343-355. The left-hand side of the internucleoside linkage is bound to the 5-membered ring as substituent  $\text{P}^*$ , whereas the right-hand side is bound to the 5'-position of a preceding monomer.

It is also clear from the above that the group P may also designate a 5'-terminal group in the case where the LNA in question is the 5'-terminal monomer. Examples of such 5'-terminal groups are hydrogen, hydroxy, optionally substituted  $\text{C}_{1-6}$ -alkyl, optionally substituted  $\text{C}_{1-6}$ -alkoxy, optionally substituted  $\text{C}_{1-6}$ -alkylcarbonyloxy, optionally substituted aryloxy, monophosphate, diphosphate, triphosphate, and  $-\text{W-A'}$ , wherein W is selected from  $-\text{O-}$ ,  $-\text{S-}$ , and  $-\text{N(R}^{\text{H}}\text{)-}$  where  $\text{R}^{\text{H}}$  is selected from hydrogen and  $\text{C}_{1-6}$ -alkyl, and where  $\text{A'}$  is selected from DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands (where the latter groups may include a spacer as defined for the substituent B).

In the present description and claims, the terms "monophosphate", "diphosphate", and "triphosphate" mean groups of the formula:  $-\text{O-P(O)}_2\text{-O}^-$ ,  $-\text{O-P(O)}_2\text{-O-P(O)}_2\text{-O}^-$ , and  $-\text{O-P(O)}_2\text{-O-P(O)}_2\text{-O-P(O)}_2\text{-O}^-$ , respectively.

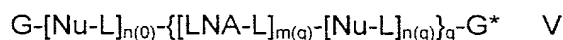


In a particularly interesting embodiment, the group P designates a 5'-terminal groups selected from monophosphate, diphosphate and triphosphate. Especially the triphosphate variant is interesting as a substrate for nucleic acid polymerases.

5

Analogously, the group P\* may designate a 3'-terminal group in the case where the LNA in question is the 3'-terminal monomer. Examples of such 3'-terminal groups are hydrogen, hydroxy, optionally substituted C<sub>1-6</sub>-alkoxy, optionally substituted C<sub>1-6</sub>-alkylcarbonyloxy, optionally substituted aryloxy, and -W-A', wherein W is selected from -O-, -S-, and  
 10 -N(R<sup>H</sup>)- where R<sup>H</sup> is selected from hydrogen and C<sub>1-6</sub>-alkyl, and where A' is selected from DNA intercalators, photochemically active groups, thermochemically active groups, chelating groups, reporter groups, and ligands (where the latter groups may include a spacer as defined for the substituent B).

15 In a preferred variant of the LNAs, the oligomer has the following formula V:



wherein

20 q is 1-50;

each of n(0), ..., n(q) is independently 0-10000;

each of m(1), ..., m(q) is independently 1-10000;

with the proviso that the sum of n(0), ..., n(q) and m(1), ..., m(q) is 2-15000;

G designates a 5'-terminal group;

25 each Nu independently designates a nucleoside selected from naturally occurring nucleosides and nucleoside analogues;

each LNA independently designates a nucleoside analogue;

each L independently designates an internucleoside linkage between two groups selected from Nu and LNA, or L together with G\* designates a 3'-terminal group; and

30 each LNA-L independently designates a nucleoside analogue of the general formula I as defined above.

Within this variant, as well as generally, the LNAs preferably include different nucleobases, in particular both nucleobases selected from thymine, cytosine and uracil and nu-

35 cleobases selected from adenine and guanine.

The oligomers are also intended to cover chimeric oligomers. "Chimeric oligomers" means two or more oligomers with monomers of different origin joined either directly or via a spacer. Illustrative examples of such oligomers which can be combined are peptides,

- 5 PNA-oligomers, oligomers containing LNA's, and oligonucleotide oligomers. The combination of an oligomer having xylo-LNA ( $R^3 = P^*$ ) domain(s) and "normal" LNA ( $R^{3*} = P^*$ ) domain(s) might be construed as an example of a chimeric oligomer as the various domains may have different affinity and specificity profiles.
- 10 Generally, the oligomers have surprisingly good hybridisation properties with respect to affinity and specificity. Thus, the oligomers comprise at least one nucleoside analogue which imparts to the oligomer a  $T_m$  with a complementary DNA oligonucleotide which is at least 2.5 °C higher, preferably at least 3.5 °C higher, in particular at least 4.0 °C higher, especially at least 5.0 °C higher, than that of the corresponding unmodified reference oligonucleotide which does not comprise any nucleoside analogue. In particular, the  $T_m$  of the oligomer is at least 2.5 x N °C higher, preferably at least 3.5 x N °C higher, in particular at least 4.0 x N °C higher, especially at least 5.0 x N °C higher, where N is the number of nucleoside analogues.
- 20 In the case of hybridisation with a complementary RNA oligonucleotide, the at least one nucleoside analogue imparts to the oligomer a  $T_m$  with the complementary DNA oligonucleotide which is at least 4.0 °C higher, preferably at least 5.0 °C higher, in particular at least 6.0 °C higher, especially at least 7.0 °C higher, than that of the corresponding unmodified reference oligonucleotide which does not comprise any nucleoside analogue. In particular, the  $T_m$  of the oligomer is at least 4.0 x N °C higher, preferably at least 5.0 x N °C higher, in particular at least 6.0 x N °C higher, especially at least 7.0 x N °C higher, where N is the number of nucleoside analogues.
- 25

The term "corresponding unmodified reference oligonucleotide" is intended to mean an oligonucleotide solely consisting of naturally occurring nucleotides which represents the same nucleobases in the same absolute order (and the same orientation).

The  $T_m$  is measured under one of the following conditions:

- 35 a) 10mM  $Na_2HPO_4$ , pH 7.0, 100mM NaCl, 0.1mM EDTA;

b) 10mM Na<sub>2</sub>HPO<sub>4</sub> pH 7.0, 0.1mM EDTA; or

c) 3M tetrametylammoniumchlorid (TMAC), 10mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.0, 0.1mM EDTA;

preferably under conditions a), at equimolar amounts (typically 1.0 μM) of the oligomer

5 and the complementary DNA oligonucleotide.

The oligomer is preferably as defined above, where the at least one nucleoside analogue has the formula I where B is a nucleobase. In a particularly interesting embodiment at least one nucleoside analogue includes a nucleobase selected from adenine and guanine.

10

Furthermore, with respect to specificity and affinity, the oligomer, when hybridised with a partially complementary DNA oligonucleotide, or a partially complementary RNA oligonucleotide, having one or more mismatches with said oligomer, should exhibit a reduction in T<sub>m</sub>, as a result of said mismatches, which is equal to or greater than the reduction which would be observed with the corresponding unmodified reference oligonucleotide which does not comprise any nucleoside analogues. Also, the oligomer should have substantially the same sensitivity of T<sub>m</sub> to the ionic strength of the hybridisation buffer as that of the corresponding unmodified reference oligonucleotide.

15

20 Oligomers defined herein are typically at least 1% modified, such as at least 2% modified, e.g. 3% modified, 4% modified, 5% modified, 6% modified, 7% modified, 8% modified, or 9% modified, at least 10% modified, such as at least 11% modified, e.g. 12% modified, 13% modified, 14% modified, or 15% modified, at least 20% modified, such as at least 30% modified, at least 50% modified, e.g. 70% modified, and in some interesting applica-  
25 tions 100% modified.

The oligomers preferably has substantially higher 3'-exonucleolytic stability than the corresponding unmodified reference oligonucleotide.

30 It should be understood that oligomers (wherein LNAs are incorporated) and LNAs as such include possible salts thereof, of which pharmaceutically acceptable salts can be relevant. Salts include acid addition salts and basic salts. Examples of acid addition salts are hydrochloride salts, sodium salts, calcium salts, potassium salts, etc.. Examples of basic salts are salts where the (remaining) counter ion is selected from alkali metals, such  
35 as sodium and potassium, alkaline earth metals, such as calcium, and ammonium ions

10

## FIGURE LEGENDS

### Figure 1.

Illustrate that only LNA primers with a matching 3'-end serve as primers in PCR.

5

### Figure 2.

Oligonucleotide-directed single base mutation obtained by incorporating 3 DNA oligos in the PCR reaction.

10 A) Starting with 3 oligonucleotides two stranded PCR primers and one mutation oligo, the mismatching base is indicated with X.

B) After 2 cycles of PCR 4 single-stranded PCR products are generated.

15 C) Product 1 and 4 generate a double-stranded PCR product containing the expected mutation after few additional PCR cycles.

### Figure 3.

Allele specific PCR with LNA primers. Left panel shows the results with primer-set

20 EQ3053/EQ3213, the "Wild type" reaction. "G" signifies that the "G-allele" of the ApoB gene was used as template, "A" that the "A-allele" of the ApoB gene was used. Right panel shows the results with primer-set EQ3157/EQ3176, the "Mutant type" reaction. Note that the primers are directed towards the non-coding strand. "C" signifies that the wildtype "G-allele" of the ApoB gene was used as template, "T" that the mutant "A-allele" of the ApoB  
25 gene was used. Marker is the Low DNA Mass™ (GibcoBRL cat. no. 10068-013, Life Technologies Inc., Rockville, MD, USA).

### Figure 4.

Comparison of LNA and DNA primers. Left panel shows the results with the LNA/DNA

30 primer-set EQ3053/EQ3213. Right panel shows the results with all-DNA primer-set EQ3647/EQ3213. "G" signifies that the "G-allele" of the ApoB gene was used as template, "A" that the "A-allele" of the ApoB gene and "H<sub>2</sub>O" that no template was included in the PCR reaction. Marker is the Low DNA Mass™ (GibcoBRL cat. no. 10068-013, Life Technologies Inc., Rockville, MD, USA).

**Figure 5.** Cycle-by-cycle monitoring of the amplification process. The increase in fluorescence during amplification of the "G-allele" (G, upper curve) and "A-allele" (A, lower curve) using the EQ3053/EQ3213 primer set and either the "A-allele" plasmid or the "G-allele" plasmid as template.

**Figure 6.** Melting curve analysis of PCR fragments generated with the EQ3053/EQ3213 primer set and either the "A-allele" template (A, lower curve) or the "G-allele" template (G, upper curve). The upper panel shows the fluorescence of the SyBR Green I dye bound to double-stranded amplicon. The lower panel shows the first negative derivative ( $-dF/dT$ ) of the melting curve in upper panel. The  $T_m$  is seen as the peak on the  $-dF/dT$  plotting.

## EXAMPLES

### Example 1

#### Sequence specific amplification with LNA primers.

##### Synthesis and analysis of primers:

Standard coupling conditions according to the protocol (0.2  $\mu$ mol scale) of the DNA-synthesizer (Pharmacia Gene Assembler Special®, Pharmacia AB, Uppsala, Sweden) were used except that the coupling time for LNA amidites was increased from the standard two minutes to five minutes. The step-wise coupling yield was approximately 99%. Standard 2'- deoxynucleoside CPG or polystyrene solid supports were generally used. After completion of the desired sequences, cleavage from the solid support and removal of protecting groups was accomplished using 32% aqueous ammonia (55°C for 5 or 10 h). LNAs synthesized in the 5'-O-DMT-ON mode were purified by reversed-phase HPLC (Delta Pak C-18, 300A, column dimensions 0.4 x 30 cm) in a concentration gradient of acetonitrile in 0.05 M triethylammonium acetate buffer pH 7.0 (flow rate 1.5 cm /min). Fractions containing 5'- O-DMT-ON LNAs (retention times of 30-35 minutes) were evaporated and de-tritylated in 80% acetic acid (1 cm, room temperature, 1 h). After evaporation, the LNAs were re-purified by reversed-phase HPLC (eluting as single peaks) as described above.

### Sample preparation

Human genomic DNA was isolated from 5 mL full blood, using the DNA Isolation Kit for Mammalian Blood (Roche Molecular Systems cat. no. 1 667 327. Roche Molecular Biochemicals, Hvidovre, Denmark) and closely following the recommendations of the manufacturer.

The "G-allele" of the ApoB gene was generated by PCR-amplification of part of the wild type human ApoB gene comprising the ApoB3500 locus (Ludwig et al. (1987) DNA 6: 363-372; accession no. M19828, SEQ ID NO 4). The PCR fragment was cloned into plasmid pCR®2.1-TOPO using the TOPO™ TA Cloning® Kit (Invitrogen, cat. no. K4500-01, Invitrogen Corporation, Carlsbad CA USA). Plasmid DNA were purified from bacterial cultures, using QUIAGEN® Plasmid Kits (QIAGEN GmbH, Hilden, Germany). The DNA sequence of the insert was verified by DNA sequencing using a ALFexpress II DNA Analysis System (Amersham Pharmacia Biotech) and closely following the recommendations of the manufacturer.

### PCR Amplification

The PCR reaction was carried out in 0.5 mL thin-wall tubes using a Hybaid PCR Express thermocycler (Hybaid, Middlesex, UK ). The dNTPs were from Amersham Pharmacia Biotech AB, Uppsala, Sweden. The final reagent mixture was composed as follows:

- 20 dATP, dGTP, dTTP, dCTP: 200 µM of each
- 1.5 mM MgCl<sub>2</sub>
- 5 µL 10 x GeneAmp PCR Buffer (Perkin-Elmer Corporation, Norwalk, CT, USA).
- 1 µM primer EQ 3053, SEQ ID NO 1
- 25 1 µM primer EQ 3054, SEQ ID NO 2
- 1 unit AmpliTaq Gold® polymerase (Perkin Elmer cat. no. N808-0240, Perkin-Elmer Corporation, Norwalk, CT, USA).
- 2 µL DNA template in serial dilutions, see figure 1.

### 30 50 µL total volume Thermocycling:

Denaturing: 95°C 15 min

Cycling: 35 times (30 sec. at 94°C; 30 sec. at 63°C; 30 sec. at 72°C)

Final extension: 10 min. 72°C

Cool down to 4°C

## 5 Detection of extension products:

10 µL of the PCR products were subsequently mixed with 2 µL application buffer (40% sucrose, 0.25% bromophenol blue, 0.25% xylene cyanol, 0.1 M EDTA pH 8.0) and electrophoresed in agarose gel. The gel was composed of 1% SeaKem®, Agarose (FMC Bio-Products, Philadelphia, PA, USA) in 1 x TAE, 0.5 µg/mL ethidium bromide (Sigma-Aldrich cat. no. E7637, Sigma-Aldrich Denmark A/S, Vallensbaek Strand, Denmark) (50 x TAE = 242 g Tris Base, 57.1 mL glacial acetic acid, 100 mL 0.5 M EDTA pH 8.0). The gels run for 0.5-1 h at approximately 5 V/cm. Fluorescence was photographed at using a Image-Master VDS equipment (Amersham Pharmacia Biotech AB, Uppsala, Sweden). A typical experiment is shown in figure 1.

To establish if 3'-terminal substitutions in PCR-primers with LNA-monomers can furnish allele specific amplification the template DNA's were amplified with the aid of the two sets of primers EQ3053/ApoBR2 and EQ3054/ApoBR2. The primer-set EQ3053/ApoBR2 (SEQ ID NO 1, SEQ ID NO 3) produce a 667 bp PCR product provided the "G-allele" of the ApoB gene (SEQ ID NO 4) is used as template, whereas the EQ3053/ApoBR2 primer-set does not produce any specific PCR product when 30.4 ng or less of the ApoB gene "A-allele" (SEQ ID NO 5) is used as template. The results are shown in figure 1.

EQ3053 (SEQ ID NO 1): 5'-CCT ACT TGA ATT CCA AGA GCA CAC **G**<sup>LNA</sup>-3'

EQ3054 (SEQ ID NO 2): 5'-CCT ACT TGA ATT CCA AGA GCA CAC **A**<sup>LNA</sup>-3'

ApoBR2 (SEQ ID NO 3): 5'-TTT AGA TCA TTT AGT TTC AGC CC-3'

"G-allele" of the ApoB gene (SEQ ID NO 4):

5'-CTTACTTGAA TTCCAAGAGC ACAC**G**GTCTT CAGTGAAGCT GCAGGGCACT-3'

"A-allele" of the ApoB gene (SEQ ID NO 5):

5'-CTTACTTGAA TTCCAAGAGC ACAC**A**GTCTT CAGTGAAGCT GCAGGGCACT-3'



## Example 2

### Detection of the ApoB R3500Q mutation by allele specific PCR with LNA-primers

Apolipoprotein B (ApoB) is the non-exchangeable protein of the low density lipoprotein

- 5 (LDL) complex that is recognised by the LDL-receptor. Familial defective ApoB 100 is an autosomal dominant disorder caused by a base transition (G -> A mutation) at amino acid 3500 (arg -> gln) - the ApoB R3500Q mutation. The mutation reduces the affinity of LDL complex for the LDL-receptor and results in elevated levels of LDL cholesterol in plasma. The prevalence rate of the mutation is 1:500 for heterozygotes and 1:1,000,000 for
- 10 homozygotes in the normal population (Harrison's Principles of Internal Medicine, 14th edition. by Fauci, A. S. et al. (Eds.) McGraw Hill ).

To demonstrate the versatility of PCR with LNA-primers two allele specific PCRs have been set up which allow positive identification of both the "G-allele" and the "A-allele" of

15 the ApoB R3500Q polymorphism.

#### Synthesis and analysis of primers:

Synthesis and analysis of LNA primers was carried out essentially as described in example 1. Downstream DNA primers were obtained as HPLC purified oligos from a commercial source (DNA Technology, Aarhus, Denmark).

#### 20 Sample preparation

The "G-allele" of the ApoB gene was generated by PCR-amplification of part of the wild type human ApoB gene comprising the ApoB3500 locus as described in example 1. Using the "G-allele" plasmid, the "A-allele" of the ApoB gene was generated by creating a oligonucleotide-directed single base mutation (G -> A mutation) at amino acid 3500 during

- 25 PCR-amplification of part of the wild type human ApoB gene comprising the ApoB3500 locus (SEQ ID NO 4). The oligonucleotide-directed single base mutation was obtained by incorporating 3 DNA oligos in the PCR reaction, two amplification oligos and one oligo covering the site to be mutated, this oligo contained one mismatch where the mutation was generated. The principle is illustrated in figure 2. Following the PCR amplification
- 30 fragments were cloned into plasmid pCR®2.1-TOPO using the TOPO™ TA Cloning® Kit (Invitrogen, cat. no. K4500-01, Invitrogen Corporation, Carlsbad CA USA). Plasmid DNA were purified from cloned bacterial cultures, using QUIAGEN® Plasmid Kits (QUIAGEN GmbH, Hilden, Germany). Plasmid preparations were then screened for presence of the expected mutation by DNA sequencing using a ALFexpress II DNA Analysis System

(Amersham Pharmacia Biotech AB, Uppsala, Sweden) and closely following the recommendations of the manufacturer.

### PCR Amplification

The PCR reactions were carried out in 0.5 mL thin-wall tubes using an Eppendorf Master-5 cyclor Gradient thermocycler (Eppendorf – Netheler – Hinz GmbH, Hamburg, Germany). The dNTP's were from Amersham Pharmacia Biotech AB, Uppsala, Sweden. The final reagent mixture was composed as follows:

### **PROCEDURE for the detection of the wildtype (G-allele) of the human ApoB3500**

#### **10 locus.**

dATP, dGTP, dTTP, dCTP: 200  $\mu$ M of each

1.5 mM MgCl<sub>2</sub>

5  $\mu$ L 10 x GeneAmp PCR Buffer (Perkin-Elmer Corporation, Norwalk, CT, USA).

15 1  $\mu$ M primer EQ 3053, SEQ ID NO 1

1  $\mu$ M primer EQ 3213, SEQ ID NO 6

1 unit AmpliTaq Gold® polymerase (Perkin Elmer cat. no. N808-0240, Perkin-Elmer Corporation, Norwalk, CT, USA).

2  $\mu$ L DNA template (approximately 50 ng plasmid/ $\mu$ L).

20

50  $\mu$ L total volume

Both the "G-allele" and the "A-allele" of the human ApoB gene generated by PCR-amplification and cloned into plasmid pCR®2.1-TOPO was used.

#### **25 Thermocycling:**

Denaturing: 95°C 15 min

Cycling: 35 times (30 sec. at 94°C; 30 sec. at 55°C; 30 sec. at 72°C)

Final extension: 10 min. 72°C

30 Cool down to 4°C

Detection:

The PCR products were subsequently analysed by standard agarose gel electrophoresis (see example 1). Only difference was 2% agarose, and stained by including GelStar® (FMC BioProducts, Rockland, Maine, USA) diluted 1:30.000 in the gel. For permanent record the gel was photographed by standard Polaroid (Polaroid LTD., St. Albans, UK) photography using an appropriate UV-transilluminator (Model TM-20E UV Products, Upland, CA, USA) and filter (Kodak Wratten #9 Eastman Kodak Co., Rochester, NY, USA).

**PROCEDURE for the detection of the mutant type (A-allele) of the human ApoB3500 locus.**

dATP, dGTP, dTTP, dCTP: 200 µM of each

1.5 mM MgCl<sub>2</sub>

5 µL 10 x GeneAmp PCR Buffer (Perkin-Elmer Corporation, Norwalk, CT, USA).

15 1 µM primer EQ 3157, SEQ ID NO 7

1 µM primer EQ 3176, SEQ ID NO 8

1 unit AmpliTaq Gold® polymerase (Perkin Elmer cat. no. N808-0240, Perkin-Elmer Corporation, Norwalk, CT, USA).

2 µL DNA template (approximately 50 ng plasmid/µL).

20

50 µL total volume

Both the "G-allele" and the "A-allele" of the human ApoB gene generated by PCR-amplification and cloned into plasmid pCR®2.1-TOPO was used

25 Thermocycling:

Denaturing: 95°C 15 min

Cycling: 35 times (30 sec. at 94°C; 30 sec. at 67°C; 30 sec. at 72°C)

Final extension: 10 min. 72°C

30 Cool down to 4°C

Detection:

The PCR products were subsequently analysed by standard agarose gel electrophoresis (see example 1). Only difference was 2% agarose, and stained by including GelStar®

(FMC BioProducts, Rockland, Maine, USA) diluted 1:30.000 in the gel. For permanent record the gel was photographed by standard Polaroid (Polaroid LTD., St. Albans, UK) photography using an appropriate UV-transilluminator (Model TM-20E UV Products, Upland, CA, USA) and filter (Kodak Wratten #9 Eastman Kodak Co., Rochester, NY, USA).

## 5 RESULTS:

To show that 3'-terminal substitutions in PCR-primers with LNA-monomers can furnish allele specific amplification template DNA's containing both the "G-allele" and the "A-allele" of the human ApoB gene were amplified with the aid of the two sets of primers EQ3053/EQ3213 and EQ 3157/EQ 3176.

10

EQ3053 (SEQ ID NO 1): 5'-CCT ACT TGA ATT CCA AGA GCA CAC **G**<sup>LNA</sup>-3'

EQ3213 (SEQ ID NO 6): 5'-GTT TTT CGT ACT GTG CTC CCA GAG-3'

EQ3157 (SEQ ID NO 7): 5'-CCC TGC AGC TTC ACT GAA GAC **T**<sup>LNA</sup>-3'

EQ3176 (SEQ ID NO 8): 5'- CAC CTC TTA CTT TTC CAT TGA GT-3'

15

"G-allele" of the ApoB gene (SEQ ID NO 4):

5'-CTTACTTGAA TTCCAAGAGC ACAC**G**GTCTT CAGTGAAGCT GCAGGGCACT-3'

"A-allele" of the ApoB gene (SEQ ID NO 5):

20 5'-CTTACTTGAA TTCCAAGAGC ACAC**A**GTCTT CAGTGAAGCT GCAGGGCACT-3'

The primer-set EQ3053/EQ3213 (SEQ ID NO 1, SEQ ID NO 6) produce a 153 bp PCR product provided the "G-allele" of the ApoB gene (SEQ ID NO 4) was used as template, whereas the primer-set does not produce any specific PCR product when the ApoB gene

25 "A-allele (SEQ ID NO 5) was used as template - see figure 3.

The primer-set EQ3157/EQ3176 (SEQ ID NO 7, SEQ ID NO 8) produce a 145 bp PCR product provided the "mutant" or "A-allele" of the ApoB gene (SEQ ID NO 5) was used as template, whereas the primer-set does not produce any specific PCR product when the

30 "wild type" or "G-allele" of the ApoB gene (SEQ ID NO 4) was used as template - see figure 3. Note that the two primer-sets EQ3053/EQ3213 and EQ3157/EQ3176 are directed towards different strands of the template molecule.

### Example 3

#### DNA primers can not discriminate between wildtype and mutant type

To evaluate the difference between LNA and DNA primers the two PCR reactions was set  
5 up.

One reaction was the reaction described in example 2 for the detection of the wildtype (G-allele) of the human ApoB3500 locus using LNA-primer EQ 3053.

- 10 The other reaction was similar but instead of using LNA-primer EQ 3053 a DNA primer with a sequence similar to EQ 3053 was used.

#### Synthesis and analysis of primers:

Synthesis and analysis of LNA primers was carried out essentially as described in example 1. DNA primers were obtained as HPLC purified oligos from a commercial source

- 15 (DNA Technology, Aarhus, Denmark).

#### PCR Amplification

The PCR reactions were carried out in 0.5 mL thin-wall tubes using an Eppendorf Mastercycler Gradient thermocycler (Eppendorf – Netheler – Hinz GmbH, Hamburg, Germany).

The dNTP's were from Amersham Pharmacia AB, Uppsala, Sweden. The final reagent

- 20 mixture was composed as follows:

The reaction with LNA-primer contained 1  $\mu$ M of primer EQ 3053 (SEQ ID NO 1) and 1  $\mu$ M of primer EQ 3213 (SEQ ID NO 6).

- 25 The reaction with DNA-primers contained 1  $\mu$ M of primer EQ 3647 (SEQ ID NO 9) and 1  $\mu$ M of primer EQ 3213 (SEQ ID NO 6).

In all other respects the reactions with LNA and DNA were identical, being:

- 30 dATP, dGTP, dTTP, dCTP: 200  $\mu$ M of each

1.5 mM  $MgCl_2$

5  $\mu$ L 10 x GeneAmp PCR Buffer (Perkin-Elmer Corporation, Norwalk, CT, USA).

1 unit AmpliTaq Gold® polymerase (Perkin Elmer cat. no. N808-0240, Perkin-Elmer Corporation, Norwalk, CT, USA).

2 µL DNA template (approximately 50 ng plasmid/µL).

5 50 µL total volume

Both the "G-allele" and the "A-allele" templates of the human ApoB gene were the plasmids discussed in Example 2.

#### Thermocycling:

- 10 Denaturing: 95°C 15 min  
 Cycling: 35 times (30 sec. at 94°C; 30 sec. at 55°C; 30 sec. at 72°C)  
 Final extension: 10 min. 72°C

Cool down to 4°C

#### 15 Detection:

The PCR products were subsequently analysed by standard agarose gel electrophoresis in 2% agarose, and stained with GelStar® (FMC BioProducts, Rockland, Maine, USA) as described in Example 2. Finally the gel was photographed as described in Example 2.

#### RESULTS:

- 20 As seen in figure 4 the primer-set EQ3053/EQ3213 (SEQ ID NO 1, SEQ ID NO 6) produced a 153 bp PCR product when the "G-allele" of the ApoB gene (SEQ ID NO 4) was used as template, whereas the primer-set did not produce any specific PCR product when the ApoB gene "A-allele" (SEQ ID NO 5) was used as template - see figure 4.

- 25 EQ3053 (SEQ ID NO 1): 5'-CCT ACT TGA ATT CCA AGA GCA CAC G<sup>LNA</sup>-3'  
 EQ3213 (SEQ ID NO 6): 5'-GTT TTT CGT ACT GTG CTC CCA GAG-3'

"G-allele" of the ApoB gene (SEQ ID NO 4):

5'-CTTACTTGAA TTCCAAGAGC ACACGGTCTT CAGTGAAGCT GCAGGGCACT-3'

30

"A-allele" of the ApoB gene (SEQ ID NO 5):

5'-CTTACTTGAA TTCCAAGAGC ACACAGTCTT CAGTGAAGCT GCAGGGCACT-3'

The all-DNA primer-set EQ3647/EQ3213 (SEQ ID NO 9, SEQ ID NO 6) however produced a 153 bp PCR product when both the "G-allele" or the "A-allele" was used as template - see figure 4.

5

EQ3647 (SEQ ID NO 9): 5'-CCT ACT TGA ATT CCA AGA GCA CAC G-3'

EQ3213 (SEQ ID NO 6): 5'-GTT TTT CGT ACT GTG CTC CCA GAG-3'

We conclude that the LNA-primer but not the similar DNA-primer can detect the single  
10 base difference of the ApoB R3500Q mutation.

#### EXAMPLE 4

##### Sequence specific amplification with LNA primer applied on the Roche Lightcycler.

To illustrate the flexibility of PCR with LNA-primers an allele specific PCR have been set  
15 up on the LightCycler (Roche Diagnostics GmbH, Mannheim, Germany). The LightCycler instrument is a device that allows amplification and identification of produced PCR products by  $T_m$  measurements in one working operation.

The reaction that allows positive identification of the "G-allele" of the human ApoB  
20 R3500Q polymorphism was as follows:

##### Synthesis and analysis of primers:

Synthesis and analysis of LNA primers was carried out essentially as described in example 1. DNA primers were obtained as HPLC purified oligos from a commercial source (DNA Technology, Aarhus, Denmark).

##### 25 PCR Amplification

The PCR reactions were carried out in the borosilicate glass capillaries provided with the instrument and closely following the recommendations of the manufacturer.

Each reaction contained:

30

- 2  $\mu$ L LightCycler - DNA Master SyBR Green I mix (Roche Diagnostics GmbH, Mannheim, Germany). The master-mix contains nucleotides, PCR buffer, SyBR Green and Taq polymerase.

- 1  $\mu$ L DNA template (approximately 50 ng plasmid/ $\mu$ L).
- 1  $\mu$ M of primer EQ 3053 (SEQ ID NO 1) and 1  $\mu$ M of primer EQ 3213 (SEQ ID NO 6).
- 3 mM  $MgCl_2$

5 In 20  $\mu$ L total volume.

Either the "G-allele" or the "A-allele" plasmids were added as template. Both the "G-allele" and the "A-allele" templates of the human ApoB gene were the plasmids discussed in Example 2.

#### 10 Thermocycling:

Denaturing: 95°C 2 min, 20°C/sec

Cycling: 44 times (0 sec. setting at 95°C; 5 sec. at 55°C; 7 sec. at 72°C)

#### T<sub>m</sub> measurement:

15 Segment 1: 95°C, 0 sec. setting.

Segment 2: 65°C, 10 sec.

Segment 3: 95°C, 0 sec. setting, 0.1°C/sec. continuous measurement.

20

Cooling: 40°C, 30 sec

#### RESULTS:

The primer-set EQ3053/EQ3213 (SEQ ID NO 1, SEQ ID NO 6) clearly amplified the "G-allele" whereas only very little PCR product was observed when the "A-allele" was used

25 as template - see figure 5.

As seen in figure 6 the primer-set EQ3053/EQ3213 produced a PCR product with a well defined T<sub>m</sub> of 82°C.

30 We conclude that the EQ3053/EQ3213 primer-set can be used for allele specific amplification in the LightCycler.